

ISSN: 0268-1242

SEMICONDUCTOR SCIENCE AND TECHNOLOGY

9th Winterschool on New Developments in Solid State Physics

Mauterndorf, Austria, 19–23 February 1996

Volume 11

Number 11S

November 1996



DTIC QUALITY INSPECTED 1

IOP Institute of Physics Publishing

A Journal Recognized by the European Physical Society



Semiconductor Science and Technology

Institute of Physics Publishing is a not-for-profit learned society publisher with a reputation for quality and high standards. It has a comprehensive range of products serving the physics and physics-related communities and in particular is an established leader in the world of scientific journals. Authors of journal articles are supplied with 25 free offprints of their articles and there are no page charges. Readers benefit from the rigorous refereeing procedures, prompt publication times and rapid response to research developments, ensuring that the journals are timely, topical and fully validated.

CoDAS Web

New alerting service featuring abstracts of condensed matter and materials science papers from the world's leading publishers in physics. Full details available from the Web: <http://www.iop.org> or via e-mail: custserv@iopublishing.co.uk.

Internet Services

For 1996, full electronic access on the World Wide Web to all Institute of Physics Publishing journals is included in the price of an institutional subscription. To gain access you will need to ask your librarian for your institution's access password. If your site is not registered your librarian (or other authorised person) will need to complete a registration form, available from our Web site (see below) or via e-mail: custserv@iopublishing.co.uk.

All physicists with access to the Internet can benefit from other information and services provided from our Web site, <http://www.iop.org>:

- *Physics Express Letters* - a subscription service to Letters and Rapid Communications from a range of our journals available prior to publication.
- CoDAS Web - condensed matter and materials science alerting service
- news about Institute of Physics activities
- complete catalogue and online books purchasing
- digests of the latest issues of our magazines
- a listing of current job vacancies in physics
- information for authors
- journal sample copy request forms

We are continually adapting and developing the services we offer to meet the needs of users. Why not use us as your home-page to ensure that you are always kept up to date?

<http://www.iop.org>



<http://www.iop.org>

Institute of Physics Publishing
Electronic Journals
- information faster than ever before

**From Spring 1996
all Institute of Physics
Publishing journals
will be available on the
World Wide Web.**

*What's more this innovative
service is included in the price of
an institutional rate subscription.*

For further information, a full list of journals and details of how to make sure your site is registered, please look at our web server on
<http://www.iop.org> or e-mail: custserv@iopublishing.co.uk

INSTITUTE OF PHYSICS PUBLISHING, TECHNO HOUSE, REDCLIFFE WAY, BRISTOL, BS1 6NX
TEL: +44 (0) 117 929 7481 FAX: +44 (0) 117 929 4318

Institute of Physics Publishing is a not-for-profit learned society publisher.

Published monthly as twelve issues per annual volume by Institute of Physics Publishing, Techno House, Redcliffe Way Bristol BS1 6NX, UK

Subscription information—1996 volume

For all countries, except the United States, Canada and Mexico, the subscription rate is £713.00 per volume. Single-issue price £59.40 (except conference issues/supplements—prices available on application). Delivery is by air-speeded mail from the United Kingdom to most overseas countries, and by airfreight and registered mail to subscribers in India.

Orders to:

Order Processing Department
Institute of Physics Publishing
Techno House, Redcliffe Way
Bristol BS1 6NX, UK.

For the United States, Canada and Mexico, the subscription rate is US\$1469.00 per volume. Delivery is by transatlantic airfreight and onward mailing.

Orders to:

American Institute of Physics, Subscriber Services
500 Sunnyside Blvd, Woodbury, NY 11797-2999, USA.

Back issues

Orders and enquiries for the previous volume should be sent to the subscription addresses given above, and for earlier volumes to:
Dawson UK Ltd, Cannon House, Folkestone CT19 5EE, UK.

United States Postal Identification Statement:

Semiconductor Science and Technology (ISSN: 0268-1242) is published monthly for US\$1469.00 per volume in association with the American Institute of Physics, 500 Sunnyside Blvd, Woodbury, NY 11797.

Postmaster:

Send address changes to:
Semiconductor Science and Technology,
American Institute of Physics, 500 Sunnyside Blvd,
Woodbury, NY 11797. Periodicals postage paid at
Woodbury, NY 11797 and additional mailing offices.

Copyright ©1996 by IOP Publishing Ltd and individual contributors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the written permission of the publisher, except as stated below. Single photocopies of single articles may be made for private study or research. Illustrations and short extracts from the text of individual contributions may be copied provided that the source is acknowledged, the permission of the authors is obtained and IOP Publishing Ltd is notified. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency under the terms of its agreement with the Committee of Vice-Chancellors and Principals. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by IOP Publishing Ltd to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of \$19.50 per copy is paid directly to CCC, 27 Congress Street, Salem, MA 01970, USA.

Printed in the UK by William Gibbons & Sons Ltd, Wolverhampton WV13 3XT.

The text of *Semiconductor Science and Technology* is printed on acid-free paper.

Semiconductor Science and Technology

Devoted exclusively to semiconductor research and applications, this journal's multidisciplinary approach reflects the far-reaching nature of this topic.

The Materials and Device Reliability section helps to link the research oriented interests of those working on physical aspects of materials and device preparation with the technical interest of those in industry looking to optimize device performance.

Honorary Editor

E Gornik, *Technical University of Vienna, Austria*

Editorial Board

G Bauer, *Johannes Kepler Universität Linz, Austria*
S P Beaumont, *University of Glasgow, UK*
C T Foxon, *University of Nottingham, UK*
E O Göbel, *Physikalisch-Technische Bundesanstalt, Braunschweig, Germany*
M Heiblum, *Weizmann Institute of Science, Rehovot, Israel*
V Kulakovskii, *Russian Academy of Sciences, Chernogolovka, Russia*
P Lugli, *Università di Roma 'Tor Vergata', Rome, Italy*
J C Maan, *University of Nijmegen, The Netherlands*
S Porowski, *Polish Academy of Science, Warsaw, Poland*
E Rosencher, *Thomson-CSF, Orsay, France*
R A Suris, *Ioffe Physico-Technical Institute, St Petersburg, Russia*
L Vina, *Universidad Autonoma, Madrid, Spain*
P Voisin, *Ecole Normale Supérieure, Paris, France*
K von Klitzing, *Max Planck-Institut für Festkörperforschung, Stuttgart, Germany*

American Sub-Board

F Capasso (Chairman), *Bell Labs, Lucent Technologies, Murray Hill, USA*
S Datta, *Purdue University, West Lafayette, USA*
S Forrest, *Princeton University, New Jersey, USA*
L Kolodziejski, *MIT, Cambridge, USA*
U K Mishra, *University of California, Santa Barbara, USA*
U Ravaioli, *University of Illinois at Urbana-Champaign, USA*
K Vahala, *California Institute of Technology, Pasadena, USA*

Japanese Sub-Board

Y Shiraki (Chairman), *University of Tokyo, Japan*
C Hamaguchi, *Osaka University, Japan*
N Sawaki, *Nagoya University, Japan*
S Tarucha, *NTT, Ibaraki, Japan*
N Yokoyama, *Fujitsu Laboratories Ltd, Atsugi, Japan*

Materials and Device Reliability Sub-Board

H E Maes (Section Editor), *IMEC, Leuven, Belgium*
M Ciappa, *ETH Zentrum, Zürich, Switzerland*
H Hartnagel, *Technische Hochschule, Darmstadt, Germany*
C Hu, *University of California, Berkeley, USA*
J W McPherson, *Texas Instruments Inc., Dallas, USA*
E Takeda, *Hitachi Ltd, Tokyo, Japan*
E Wolfgang, *Siemens AG, Munich, Germany*

Editorial
S Counsell

Production
A D Evans

Publishing Office
Institute of Physics Publishing
Techno House, Redcliffe Way
Bristol BS1 6NX, UK
Tel: +44 (0)117 929 7481
Fax: +44 (0)117 929 4318
Internet: prod1@ioppublishing.co.uk

Consultant Editor
V Grigor'yants
IOPP Editorial Office
Ioffe Physico-Technical Institute
26 Polytechnicheskaya
194021 St Petersburg, Russia
Advertisement Sales
Jack Pedersen/John Irish
D A Goodall Ltd
17 Blossom Street
London E1 6JH
Tel: +44 (0)171 375 1155
Fax: +44 (0)171 375 1115

1 999 0222 062

Semiconductor Science and Technology

Journal scope

Experimental and theoretical studies of the structural, electrical, optical and acoustic properties and the doping of bulk, low-dimensional and amorphous semiconductors; computational semiconductor physics; interface properties, including the physics and chemistry of heterojunctions, metal-semiconductor and insulator-semiconductor junctions; all multi-layered structures involving semiconductor components.

Physics of semiconductor devices, including theoretical modelling and experimental demonstration; all aspects of the technology of semiconductor device and circuit fabrication; growth and preparation of materials, including both epitaxial (e.g. molecular beam and chemical vapour methods) and bulk techniques; the structural, electrical and optical characterization of material and device structures.

Also included are appropriate aspects of surface science; the influence of growth kinetics and chemical processing on layer and device properties; growth models; relevant areas of 'molecular electronics' and semiconductor structures incorporating Langmuir-Blodgett films; resists, lithography and metallization where they are concerned with the definition of small geometry structure.

Articles in *Semiconductor Science and Technology* are abstracted in: INSPEC; FIZ; Bulletin Signaletique; AEA (UK INIS/IEA Energy Technology Data Exchange); Chemical Abstracts; Engineering Information; INIS; ISI (Science Citation Index, Scisearch, Research Alert; Materials Science Citation Index, Current Contents/Physical, Chemical and Earth Sciences, Current Contents/Engineering, Technology and Applied Sciences).

A brief guide for authors

A submission to *Semiconductor Science and Technology* must be the original work of the author(s) and must not be published elsewhere or under consideration for another publication in its submitted or a substantially similar form in any language. Research Papers (up to 8500 words) and Letters to the Editor (up to 2500 words) will be considered. They may be in English, French or German, but an abstract, title and list of figure and table captions in English must be provided. The following material should be submitted.

- (1) Three copies of the typescript, at least one of which must be single-sided, each including:
 - (a) Title page with title of article, name(s) of author(s) and address(es) of establishment(s) where work was carried out.
 - (b) Abstract.
 - (c) Text with double line spacing on A4 (210 mm x 297 mm) or similar paper.
 - (d) List of references given in either the numerical or the Harvard (alphabetical) system (on separate sheet).
 - (e) List of captions for illustrations.
 - (f) Set of copies of illustrations.
- (2) One complete set of illustrations suitable for reproduction. The quality of labelling, line thickness and other detail should be adequate to maintain acceptable size, density and clarity after an appropriate reduction in size. Photographic illustrations should be supplied as glossy prints (not negatives or slides). Graphics files may also be supplied (see below) but good quality originals are still required in case the electronic files cannot be used.
- (3) Copies of any unpublished or obscure references that may be necessary for the refereeing process.
- (4) A completed Assignment of Copyright form. Authors unable to include this with their submissions will be sent a copy for signature.

Colour reproduction

Colour reproduction of illustrations is available in four colours or spot colour, for which the author will be asked to pay the additional reproduction costs incurred.

Address for submissions

The typescripts and figures should be sent to:

Managing Editor
Semiconductor Science and Technology
IOP Publishing Ltd
Techno House
Redcliffe Way
Bristol BS1 6NX UK

Articles in electronic formats

Articles can be published from a wide range of electronic text (TeX (including LaTeX, REVTeX, AMS-TeX, etc) source code, wordprocessor (e.g., Word, WordPerfect) files or ASCII files) and graphics formats. Authors are asked to submit hard-copy versions for refereeing as normal and to indicate in their submission letter the formats in which the text and illustrations are available. Further details are available from Institute of Physics Publishing's World Wide Web server (<http://www.ioppublishing.com>) and in the booklet Notes for Authors (see below).

Further details

Authors who are submitting to *Semiconductor Science and Technology* for the first time, or who require more details on presentation and style, should consult the booklet Notes for Authors, obtainable free of charge from the Journals Publishing Department at the Publishing Office. Email requests should be addressed to notes4au@ioppublishing.co.uk. Notes for Authors is also available in electronic format via the World Wide Web server (<http://www.ioppublishing.com>)

Foreword

The Ninth International Winterschool on New Developments in Solid State Physics, entitled *Nanostructure Physics and Technology*, was held in Mauterndorf Castle, Salzburg, Austria from 19 to 23 February 1996. A total of 65 papers (including posters) were presented at the meeting. This issue contains 30 invited papers.

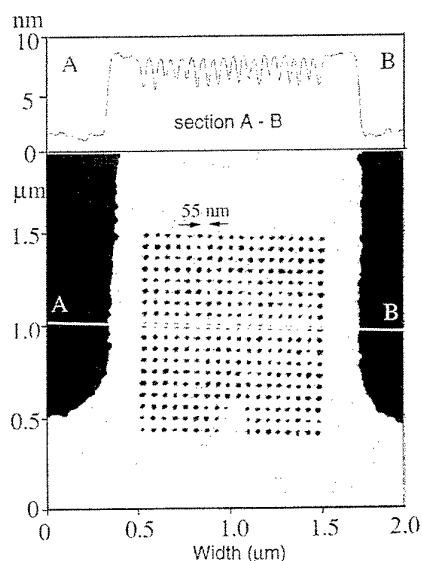
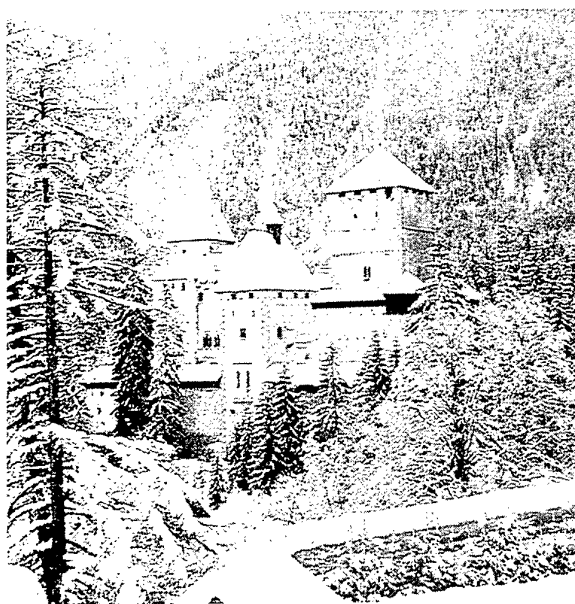
Following the tradition of the previous meetings, a serious attempt was made to have the most recent highlights in low dimensional physics as well as advances in nanotechnology presented at this meeting. The main topics were:

- Composite fermions and skyrmions
- Single-electron phenomena
- Quantum dots
- Magnetotransport and quantum Hall effect
- Nanolithography, nanofabrication and nanoprobe
- Superlattices and multiple quantum wells
- Mesoscopics

The success of this conference series relies heavily on the invited speakers, whose efforts resulted in lucid and compelling presentations of their work. The event's strong international tradition was maintained by a total of about 170 scientists attending from 13 countries.

A large number of people contributed through their advice, support and collaboration to the conference: to all of them the organizers are most grateful.

G Bauer, F Kuchar and H Heinrich
Conference Co-Editors



9th International Winterschool on New Developments in Solid State Physics

NANOSTRUCTURE PHYSICS AND TECHNOLOGY

19–23 February, 1996
Mauterndorf, Salzburg, Austria

Organizing Committee

F Kuchar Montanuniversität Leoben, Austria
G Bauer Johannes Kepler Universität Linz, Austria
H Heinrich Johannes Kepler Universität Linz, Austria

Local Organization

H Kirchberger
S Lechner
K Rabeder
H Sitter

The conference organizers would also like to thank the following companies for their support:

Balzers-Pfeiffer AG, Austria
Bruker Analytische Meßtechnik, Germany
Cryophysics, Germany
EG & G, Germany
IBM, Austria
Dr Müller Vakuum-Dünnschicht-Technik, Germany
Omicron, Germany
Oxford Instruments, Germany

The conference acknowledges substantial support by the Bundesministerium für Wissenschaft, Forschung und Kunst, Austria, Österreichische Physikalische Gesellschaft, Austria, Österreichische Forschungsgemeinschaft, Austria, Gesellschaft für Mikroelektronik, Austria, US Air Force EOARD, London, UK, Office of Naval Research, USA

This work relates to Department of Navy Grant N00014-96-1-0513 issued by the office of Naval Research. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein.

Examining composite fermions with surface acoustic waves

R L Willett and L N Pfeiffer

AT&T Bell Laboratories, 600 Mountain Avenue, Room 1D-466, Murray Hill, NJ 07974, USA

Abstract. Fermi surface effects of composite fermions (CF) in a two-dimensional electron system are exposed by using large-wavevector surface acoustic waves (SAW) to probe the system conductivity. Enhanced conductivity is observed at even-denominator filling factors for SAW wavelengths smaller than the CF mean free path. Larger-wavevector SAW demonstrate geometric resonance with the CF cyclotron orbit, allowing direct measurement of the CF Fermi wavevector. Emphasized here are recent geometric resonance observations which display an apparent inconsistency between the effective mass as measured in d.c. transport experiments and the mass necessary to resolve the geometric resonances. Using these same high-frequency SAW, a geometric resonance structure at $\nu = 3/2$ defines a composite fermion Fermi wavevector consistent with a spin-polarized CF population: this is at odds with d.c. transport measurements. Present efforts at determining composite fermion structure using large-wavevector SAW will also be discussed.

Surface acoustic waves (SAW) can be used to examine the 2D electron system (2DES) response at small length scales, that of the SAW wavelength. Our experiments employ SAW propagated on GaAs/AlGaAs heterostructures to extract the 2D electron wavevector-dependent conductivity $\sigma_{xx}(q)$. The SAW traverses the 2DES where it is attenuated and slowed by the interaction of the piezoelectric field due to the ultrasound and the 2D electrons' ohmic response. The sound velocity shift as measured here decreases monotonically with increasing 2DES conductivity: $\Delta v/v = \alpha/[1 + (\sigma_{xx}(q)/\sigma_m)^2]$, with the effective piezoelectric coupling $\alpha = 0.00032$, and with σ_m used as a scaling parameter. The essential point of the experiments is that the measured amplitude and velocity reflect the 2DES conductivity at the dimension of the SAW wavelength. As such this technique has been used [1] to study the integral quantum Hall effect, the fractional quantum Hall effect, the small filling factor regime, and is now also being employed as a contactless probe of one-dimensional and zero-dimensional systems. Our fundamental technological goal has been to develop larger-wavevector SAW, and we have produced new SAW that operate up to 10.7 GHz, which allow examination of the 2DES conductivity at a length scale of ~ 2700 Å.

Measuring the 2DES response at small length scales has been crucial in exposing the properties of the gauge transformed or composite fermion [2]. In particular, surface acoustic waves have revealed the composite fermion's Fermi surface effects at even-denominator filling factors as delineated by Halperin *et al* [3] (HLR).

At low SAW wavevectors and frequencies, the measured sound velocity shift describes a sheet conductivity throughout the magnetic field range essentially similar to

that seen in d.c. transport. However, in our early results [4], SAW demonstrated an anomalous drop in transmitted amplitude and velocity at $\nu = 1/2$ and $1/4$, reflecting an enhanced conductivity of the 2DES at these specific ν (see figure 1). Following HLR, this anomaly is understood to occur when the SAW wavelength is smaller than or about the dimension of the composite fermion mean free path, so that the conductivity is enhanced since the quasiparticle can move in the direction of the piezoelectric field of the SAW without scattering. The width of the anomaly in a magnetic field is determined by the cyclotron motion of the quasiparticle: the enhanced conductivity is cut off because the quasiparticle moves more laterally with respect to the SAW propagation direction as the effective B -field, $B_{\text{eff}} = B - B(1/2)$, increases. Further, the theory proposed that the conductivity at $\nu = 1/2$ should increase linearly with SAW wavevector as should the width of the anomaly: these predictions were found to be true experimentally [5].

Beyond these findings, predicted definitive evidence for the presence of a Fermi surface at $1/2$ is observation of geometric resonance of the quasiparticle cyclotron motion and the SAW wavelengths; these resonances provide a direct measurement of the composite fermion Fermi wavevector, k_F . Within the above described enhanced conductivity, resonance structure should appear in the conductivity ($\Delta v/v$) if composite fermion cyclotron orbits can be completed within the SAW wavelength without scattering. By using smaller-wavelength SAW, such resonances were indeed observed [6] in SAW measurements up to 8.5 GHz. In figure 1 these resonances are shown to develop symmetrically about $1/2$ as larger SAW wavevector is used, with the resonance positions in ΔB increasing linearly with SAW q : $\Delta B/B = q/3.83k_F$ for the

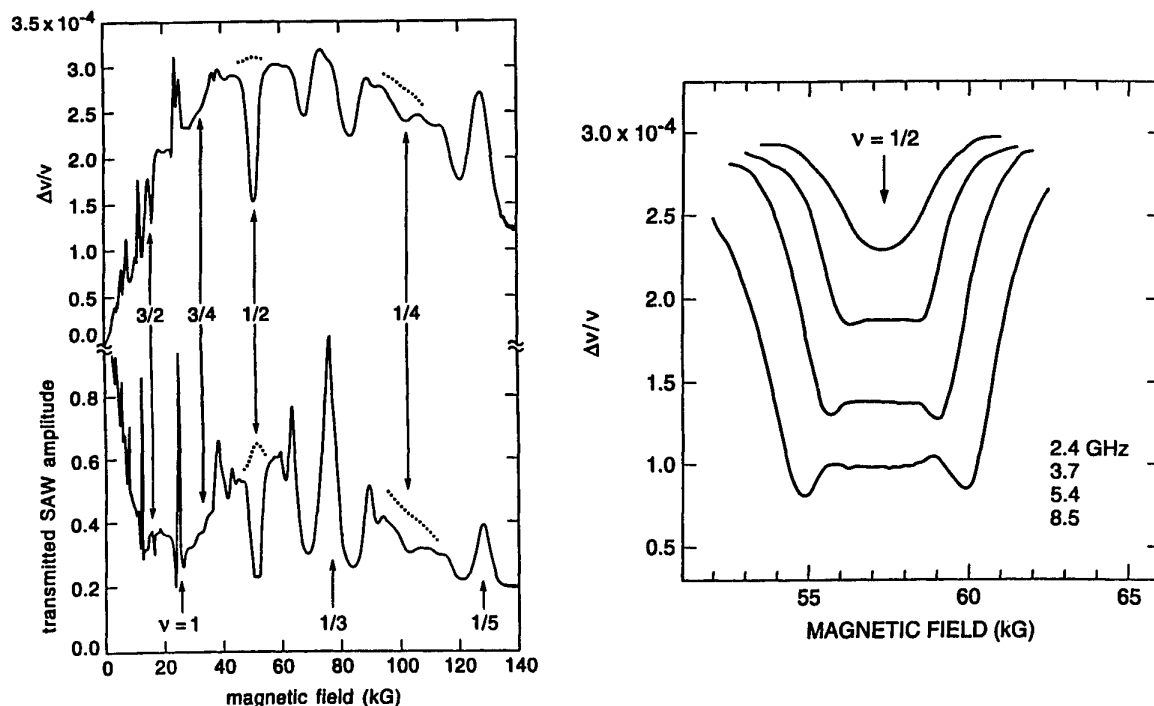


Figure 1. Left panel: SAW velocity shift and transmitted amplitude for a 3.4 GHz SAW at 120 mK versus magnetic field in wafer A. The dotted curves represent $\Delta v/v$ and amplitude calculated from the d.c. conductivity. Right panel: sound velocity shift versus B -field at four SAW frequencies demonstrating geometric resonances developing symmetrically about $\nu = 1/2$ in wafer B.

principal resonance positions, corresponding to resonance of the CF cyclotron orbit with a single SAW wavelength. This commensurability of the cyclotron orbits and the wavelength allows measurement of the quasiparticle k_F , which closely matches [6] the value prescribed by the electron system real density n , with $k_F = (4\pi n)^{1/2}$.

We have further developed the SAW technology and have studied the 2DES conductivity up to 10.7 GHz, corresponding to a wavelength of ~ 2700 Å. These results demonstrate geometric resonance structure at multiple even-denominator filling factors, show both secondary and principal resonances at $\nu = 1/2$, and most importantly reveal a fundamental inconsistency with the effective mass measurements [7] using simple d.c. transport.

The higher sample density and smaller SAW wavelengths allow observation of both principal and secondary resonances around $1/2$, corresponding to commensurability of the cyclotron orbits with both one and two SAW wavelengths (see figure 2). The resonance structure agrees with the theoretical curve using experimentally derived values of quasiparticle mean free path and electron density inhomogeneity (see [7]). The conditions necessary for resonance observation place constraints on the quasiparticle cyclotron frequency, and therefore the mass. In order to observe a resonance, the SAW must present to the quasiparticle orbits a static spatial wave, similar to a grating; quantitatively this is equivalent to $\omega_c^* \gg \omega_{\text{SAW}}$, and $v_f^* \gg v_{\text{SAW}}$. Given observation of the resonances with high resolution, we then compare these satisfied constraints to

the effective mass as derived from d.c. transport studies on these same samples. By measuring activation energies and Shubnikov-de Haas amplitudes of the fractional quantum Hall effect (FQHE) states as done previously [8], we extract cyclotron frequency and effective mass values over a magnetic field range near the SAW resonance positions (see figure 2, right panel). The resonance positions in the B -field are marked on the same scale showing the SAW frequency and the composite fermion cyclotron frequencies for masses m_e and $1.4m_e$. This corresponds to the range of masses derived from d.c. transport (shown to the left). We find that these mass values are too large to be consistent with observation of the SAW resonances. As shown in the figure, when a quasiparticle mass of $1.4m_e$ is used, both the cyclotron frequencies at the principal and secondary resonance positions are less than or nearly equal to the SAW frequency, thus violating the static wave constant of $\omega_c^* \gg \omega_{\text{SAW}}$.

The origin of the inconsistency is unclear. Three possibilities exist: the d.c. transport derived effective mass is wrong, the conditions considered above to be necessary for resonance observation are incorrect, or a different cyclotron frequency is applicable for the d.c. transport and the SAW. As such, these results provide a new circumstance for theoretical evaluation of the composite fermion dynamics over the magnetic field spectrum.

Geometric resonances can be observed at multiple even-denominator filling factors using large-wavevector SAW. As was the case for $\nu = 1/2$, resonance structure appear

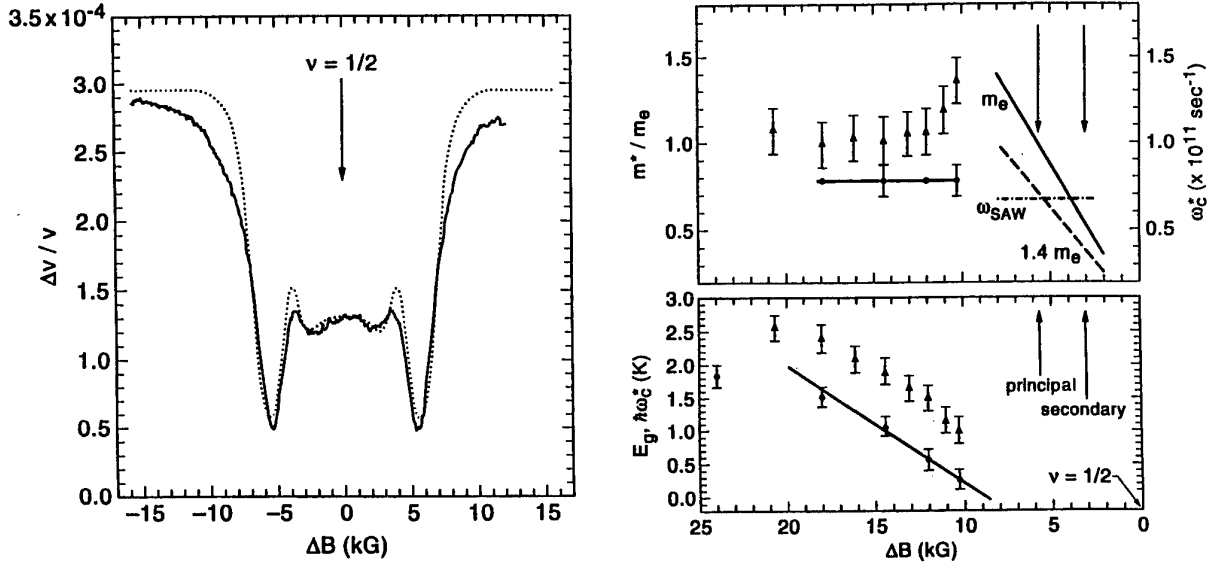


Figure 2. Left panel: sound velocity shift versus effective B -field for a 10.7 GHz SAW around $\nu = 1/2$ demonstrating both principal and secondary geometric resonances. The dotted curve is a theoretical fit as outlined in [7]. Right panel: energy gap, cyclotron energy and effective mass as derived from d.c. transport. The cyclotron frequency using $m^* = 1.0m_e$ and $1.4m_e$, with m_e the electron mass, is compared with the SAW frequency in the upper right. The vertical lines mark the B -field positions of the geometric resonances shown in the left panel.

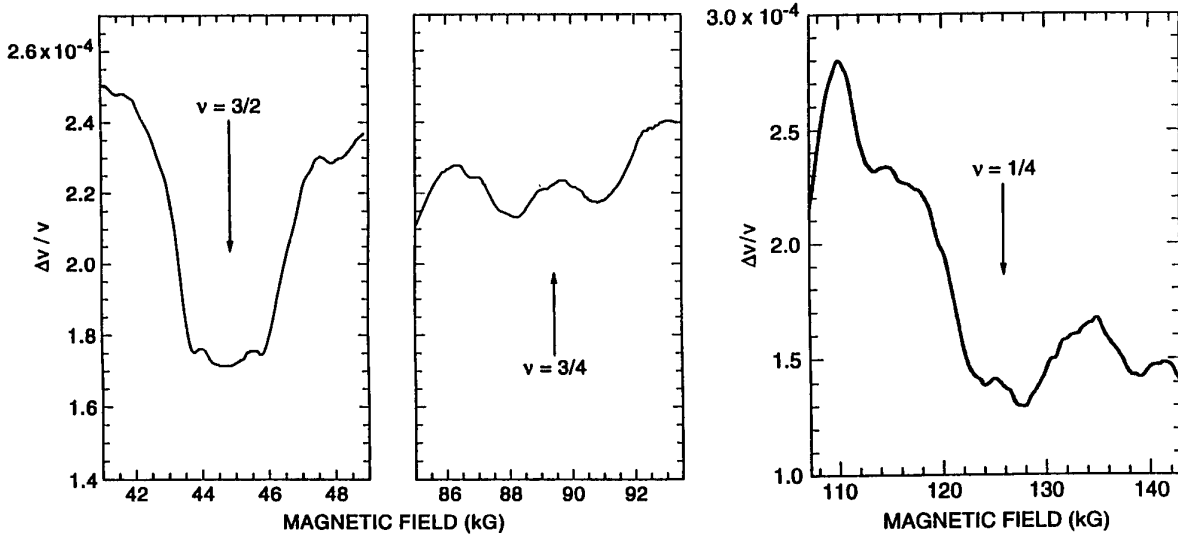


Figure 3. Geometric resonance structure around filling factors $3/2$, $3/4$ and $1/4$. The data for $3/2$ and $3/4$ are from a SAW frequency of 10.7 GHz. The data in the right panel for $\nu = 1/4$ are from an 8.5 GHz SAW using a different wafer. Temperatures in both are about 130 mK.

within the enhanced conductivity feature (minimum in $\Delta v/v$) for filling factors $3/2$, $3/4$ and $1/4$ (see figure 3). The positions in magnetic field of the resonances at $3/2$ provide a measure of the Fermi wavevector for this filling factor: the measured k_F corresponds to a filled lower spin state and a half-filled upper energy spin state. The accuracy of this determination of the spin state filling is limited by the accuracy with which the resonance position can be assigned in conjunction with the weak dependence of Fermi

wavevector on density, $k_F = (4\pi n)^{1/2}$. The resonance positions can only be determined to about an accuracy of 10% from the data shown, which translates to an accuracy in density determination of about 20%. Therefore, the data show that the lower spin state is filled and the upper spin state is empty within an error that could allow up to a mixed spin state population of about 20%. If indeed the lower spin state is fully populated, this is contrary to recent d.c. transport measurements [9] that indicate a mixed spin

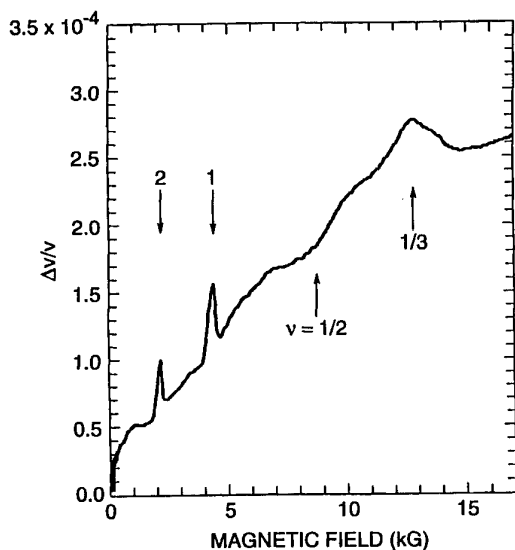


Figure 4. SAW velocity shift for a 900 MHz SAW at 300 mK in an ultra-low-density heterostructure, clearly demonstrating composite fermion Fermi surface effects at 0.8 T and the FQHE at 1.2 T.

configuration at $\nu = 3/2$. The aetiology of this discrepancy is not known. The enhanced conductivity at filling factors $3/4$ and $1/4$ (figure 3) is resolved into broad resonance structures with peaks occurring at effective magnetic fields predicted by theory and consistent with the resonance positions at $\nu = 1/2$. These resonances are notably broader than both the $1/2$ and $3/2$ structures and imply that the Chern–Simons gauge transformation employing a greater number of flux quanta is more susceptible to density variations.

The results of these experiments and other investigations elicit some of the major questions related to our understanding of these new quasiparticles and their interactions. Two important concerns in this picture of the fermion Chern–Simons construction are (i) what is the effective mass of the quasiparticle and what determines this mass, and (ii) what is the structure of the composite fermion, or more specifically, at what point does a mean-field construction break down. The latter question relates to the nature of scattering by the gauge field, which experimentally is thus far the principal differentiation between composite fermion

dynamics and those of the electron. In an attempt to answer these issues, and as progressively larger-wavevector SAW have revealed more properties of these quasiparticles, it is reasonable to attempt to push the length scale of the SAW toward the fundamental length scale of the 2D system, its Fermi wavevector. The SAW experiments performed to date have all operated in the hydrodynamic limit, with $k_F \gg k_{\text{SAW}}$. If it is technologically possible to move into the regime of $k_{\text{SAW}} > k_F$, then the SAW method may in effect become a structural probe. While we have succeeded in increasing the SAW wavevector substantially, figure 4 demonstrates another step toward reaching this structural probe regime. The figure shows relatively low-wavevector SAW measurements of an ultra-low-density heterostructure. Note the features of the integral quantum Hall effect at $\nu = 1$ and 2 and the clear response at filling factor $1/3$. Remarkably, enhanced conductivity at $\nu = 1/2$ is also apparent, occurring at a magnetic field of only 0.8 T. This low density corresponds to a composite fermion Fermi wavevector of about $35 \mu\text{m}^{-1}$; a 10 GHz SAW has a wavevector of $22 \mu\text{m}^{-1}$. With further development of high-quality, low-density heterostructures, and advances in SAW technology, structural studies of the correlated electron system may become a possibility.

References

- [1] See Wixforth A, Kotthaus J P and Weimann G 1986 *Phys. Rev. Lett.* **56** 2104
Paalanen M A, Willett R L, Littlewood P B, Ruel R R, West K W, Pfeiffer L N and Bishop D J 1992 *Phys. Rev. B* **45** 11 342
- [2] Jain J K 1990 *Phys. Rev. Lett.* **63** 199
- [3] Halperin B I, Lee P A and Read N 1993 *Phys. Rev. B* **47** 7312
- [4] Willett R L, Paalanen M A, Ruel R R, West K W, Pfeiffer L N and Bishop D J 1990 *Phys. Rev. Lett.* **65** 112
- [5] Willett R L, Ruel R R, Paalanen M A, West K W and Pfeiffer L N 1993 *Phys. Rev. B* **47** 7344
- [6] Willett R L, Ruel R R, West K W and Pfeiffer L N 1993 *Phys. Rev. Lett.* **71** 3846
- [7] Willett R L, West K W and Pfeiffer L N 1995 *Phys. Rev. Lett.* **75** 2988
- [8] Du R R *et al* 1993 *Phys. Rev. Lett.* **70** 2944; 1994 *Phys. Rev. Lett.* **73** 3274
Leadley D R *et al* 1994 *Phys. Rev. Lett.* **72** 1906
- [9] Du R R, Yeh A S, Stormer H L, Tsui D C, Pfeiffer L N and West K W 1995 *Phys. Rev. Lett.* **75** 3926
See also Barrett S E, Dabbagh G, Pfeiffer L N, West K W and Tycko R 1995 *Phys. Rev. Lett.* **74** 5112

The dependence of the Composite Fermion effective mass on carrier density and Zeeman energy

R J Nicholas[†], D R Leadley[‡], M S Daly[†], M van der Burgt[†],
P Gee[†], J Singleton[†], D K Maude[§], J C Portal[§], J J Harris^{||} and
C T Foxon[¶]

[†] Department of Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK

[‡] Department of Physics, Warwick University, Coventry CV4 7AL, UK

[§] Laboratoire des Champs Magnétiques Intense, CNRS, F38042 Grenoble, Cedex 9, France

^{||} Department of Electronics and Electrical Engineering, University College London, UK

[¶] Department of Physics, University Park, Nottingham NG7 2RD, UK

Abstract. Measurements of the temperature-dependent resistivity of high-mobility GaAs/GaAlAs heterojunctions are used to measure the effective mass of Composite Fermions (CF). The CF effective mass is found to increase approximately linearly with the effective field B^* up to effective fields of 14 T. Data from all fractions around $\nu = 1/2$ are unified by the single parameter B^* for samples studied over a wide range of temperature. The energy gap is found to increase as $\sqrt{B^*}$ at high fields. Hydrostatic pressure is used to reduce the value of the electron g -factor, and this is shown to have a large effect on the relative strengths of different fractions. By 13.4 kbar, where the Zeeman energy is only 1/4 of its value at 0 bar, fractions with odd numerators are found to be strongly suppressed, and new features with even numerators appear. The energy gaps measured for 5/3 as a function of carrier density and pressure are consistent with a g -factor equal to the bulk value enhanced by a factor of two due to exchange interactions.

The fractional quantum Hall effect has been known for many years [1, 2], but our picture of the phenomenon is changing rapidly due to the introduction of the Composite Fermion (CF) model [3, 4]. In this model the Coulomb interaction of one electron with all the others is replaced with a Chern–Simons gauge field, equivalent to attaching an even number ($2m$) of flux quanta ($\Phi_0 = h/e$) to each electron. In a mean field approximation the gauge field exactly balances the external field at filling factor $\nu = 1/2m$ where the system of interacting electrons in high magnetic field is replaced by one of independent CFs in zero field. At other filling factors there are more (or fewer) flux quanta than required to cancel the gauge field and the CFs see an effective magnetic field, $B^* = B - 2m\Phi_0 n_e$. This leads to quantization of the CF energy into Landau levels (LLs) and gaps open in exact analogy with the integer quantum Hall effect (IQHE) of non-interacting electrons. Thus the FQHE (fractional QHE) may be simply regarded as the IQHE of composite fermions. This behaviour is illustrated in figure 1, which shows the resistivity of a high-mobility GaAs/GaAlAs heterojunction as a function of the magnetic field, written in terms of the density of flux quanta per electron. In the region $0 < B/n_e < 1$ we see the Shubnikov–de Haas oscillations of the single-

particle electrons, while from $1 < B/n_e < 3$, corresponding to $-1 < B^*/n_e < 1$, we see the Shubnikov–de Haas oscillations of the first generation ($m = 1$) composite fermions. From $3 < B/n_e < 4$, corresponding to $-1 < B^*/n_e < 0$, we see second-generation ($m = 2$) oscillations. This picture becomes more complex in the regions of $0.5 < B, B^*/n_e < 1$ when regarded in detail, since further features appear at fractions such as 4/3 and 5/3 and their higher-generation analogues 5/7 and 4/5. In this case mixed states appear which consist of filled levels of more than one type of particle, such as single-particle electrons in a filled Landau level and CF states formed from the remaining electrons.

By treating the FQHE features in ρ_{xx} as CF Shubnikov–de Haas oscillations, it is possible to deduce an effective mass of the CF particles [5–8] by analysing the temperature dependence of the oscillation amplitudes [9]. Work on low-density samples suggested that the effective mass M^* could be described by the relation $M^* = 0.51 + (0.074)B^*$, in units of the free electron mass m_e . The same dependence was found for positive and negative values of B^* , and the mass values from samples with n_e differing by a factor of two was found to be the same at a given value of B^* . There have also been reports that the mass may diverge as

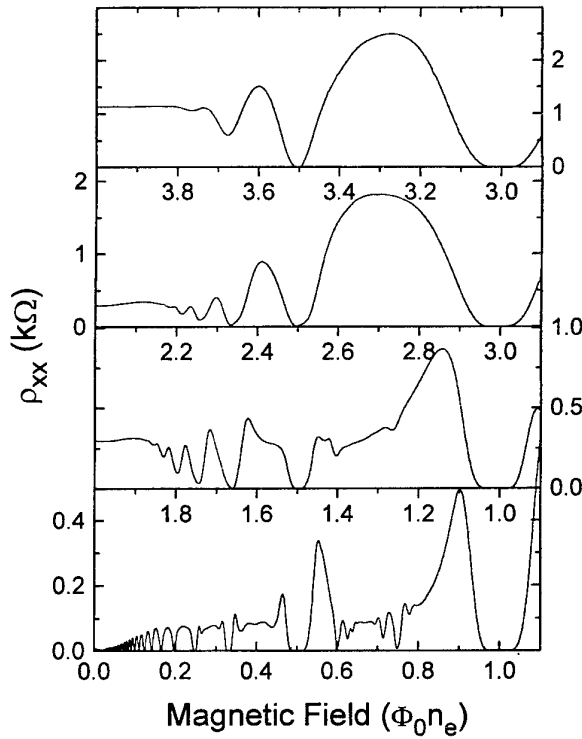


Figure 1. A plot of the FQHE features in a GaAs/AlGaAs heterojunction with a carrier density of $9 \times 10^{10} \text{ cm}^{-2}$ measured at 97 mK. The data have been replotted as a function of magnetic field to emphasize the symmetry of the oscillations.

$\nu = 1/2$ is approached [6–8], where the oscillations in ρ_{xx} are weak. Studies of the mass have recently been extended to much higher densities by the use of pulsed magnetic fields up to 50 T, combined with ^3He temperatures [10]. In these measurements good thermal equilibrium was ensured by immersing the sample in liquid, and using low measuring currents and keeping induced voltages to a minimum during the 10 ms duration pulses. Some typical experimental recordings are shown in figure 2, for samples with carrier densities of 3.2 and $4.8 \times 10^{11} \text{ cm}^{-2}$. The oscillations in resistivity are analysed using the Ando formula [9].

$$\frac{\Delta\rho_{xx}}{\rho_0} \propto \frac{X}{\sinh X} \exp\left(-\frac{\pi}{\omega_c \tau_q}\right) \cos 2\pi\left(\nu - \frac{1}{2}\right) \quad (1)$$

where $X = 2\pi^2 k_B T / \hbar \omega_c$ and $\omega_c = eB/m^*$ is the cyclotron frequency. For composite fermions we replace B by B^* , ν by ν^* , m^* by M^* and τ_q by T_q . The analysis was performed on the features that are large at low temperature for the range of conditions corresponding to $\Delta\rho/\rho < 0.5$, where the oscillations are only a weak modulation of the conductivity and the higher harmonics can be neglected.

The results of the analysis are shown in figure 3 as a function of n_e for several different fractions. There is no unique n_e dependence covering all fractions, but it can clearly be seen that instead the mass values fall in pairs, corresponding to states with a common numerator p , for example $2/3$ and $2/5$. These have equal numbers of occupied CF Landau levels, but occur on either side

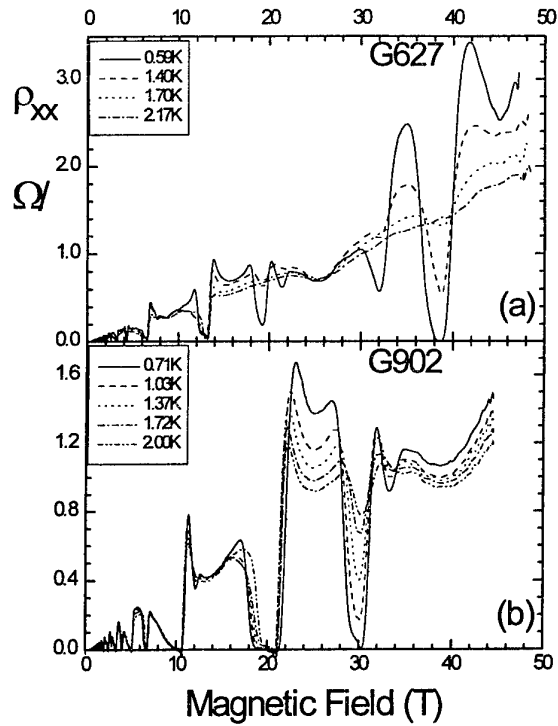


Figure 2. Typical experimental traces taken in two GaAs/AlGaAs heterojunctions with carrier densities of 3.0 and $4.8 \times 10^{11} \text{ cm}^{-2}$ measured in a pulsed magnetic field as a function of temperature.

of $\nu = 1/2$ with effective fields in opposite senses. This provides a simple demonstration of the symmetry of the states about $\nu = 1/2$ which is consistent with the CF model, rather than that of particle-hole conjugation where states of common denominator q (for example $1/3$ and $2/3$) look similar. In the low-density limit all states tend to the same effective mass, but the lower-index CF Landau levels show an increasing ‘non-parabolicity’. By $n_e = 3 \times 10^{11} \text{ cm}^{-2}$ the effective masses for $1/3$ and $2/3$ differ by approximately 40%. In addition, the gradients of each line in figure 3 show an accurate ($\pm 2\%$) $1/p$ dependence. The CF mass may therefore be described by the expression

$$M^* = 0.51 + \frac{0.35}{p} n_e = 0.51 + 0.083 B^* \quad (2)$$

in units of m_e , with n_e in units of 10^{11} cm^{-2} .

The dependence of M^* on B^* is shown in figure 4(a) for a wide range of samples, which shows that there is a simple functional dependence of M^* on the single parameter B^* , covering more than a factor of 25 variation in B^* . The measurement of M^* is equally a measurement of the CF cyclotron energy $E_c^* = \hbar e B^* / M^*$. This E_c^* is equivalent to what would previously have been known as the FQHE energy gap Δ , for a sample with infinitely narrow levels. In figure 4(b) we show E_c^* as a function of B^* for all of the samples studied. The increase of M^* with field necessarily means that E_c^* shows a sublinear increase with effective field. The broken curve in figure 4(b) is a fit to the data with $E_c^* = a\sqrt{B^*}$, where $a = 3.3 \text{ K T}^{-1/2}$.

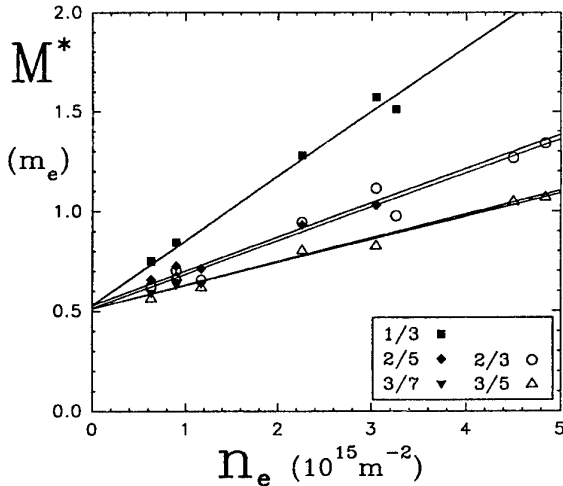


Figure 3. The CF effective mass as a function of carrier density for different fractions. The full lines show least squares fits.

Although this is not the function generated by a mass that increases linearly from an offset zero, above a value of $B^* = 1$ T there is remarkably good agreement with this single dependence on B^* alone.

The FQHE is the result of a many-body Coulomb interaction, and so theoretically we would expect the energy gaps to scale with a relation of the form $\Delta = C_v e^2 / (4\pi\epsilon\epsilon_0 l_0)$, where l_0 is the cyclotron radius (proportional to the interparticle spacing) and C_v is a fixed coefficient, different for each fraction. Halperin *et al* [4] have used this relationship and argued on dimensional grounds that the high-field limit of M^* should show a \sqrt{B} dependence with changing carrier density through l_0 . This corresponds to a \sqrt{B} dependence for both E_c^* and M^* for any given fraction, but our results suggest that there is not a single functional dependence on n_e but instead on B^* .

For fractions at values of occupancy greater than 1 the picture is more complex, since both spin states of the lowest single-particle Landau level must have some finite occupancy. It has long been known that increasing the value of the spin splitting by tilting the sample relative to the applied magnetic field can influence the relative strength of the FQHE states [11–13]. Recently Du *et al* [14] have shown that for fractions greater than 1 it is possible to analyse the CF states as the oscillations due to CF holes with a concentration $N_h^* = 2B/\Phi_0 - n_e$ in an effective field $B^* = 2n_e\Phi_0 - 3B$. In this picture the 5/3 state consists of a single occupied hole CF Landau level, with its gap determined by the smaller of E_c^* and the Zeeman energy $g^*\mu_B B$, while 4/3 has two occupied levels, which for the untilted case represents an unpolarized state with one level from each spin. On increasing the Zeeman energy by tilting, the two families of CF Landau levels move through each other causing features with odd and even numbers of filled CF levels to oscillate in strength. This led them to deduce a value for g^* of $0.61 + 0.087B^*$ for one density.

In this paper we have varied both the carrier density and studied the effects of *decreasing* the Zeeman energy

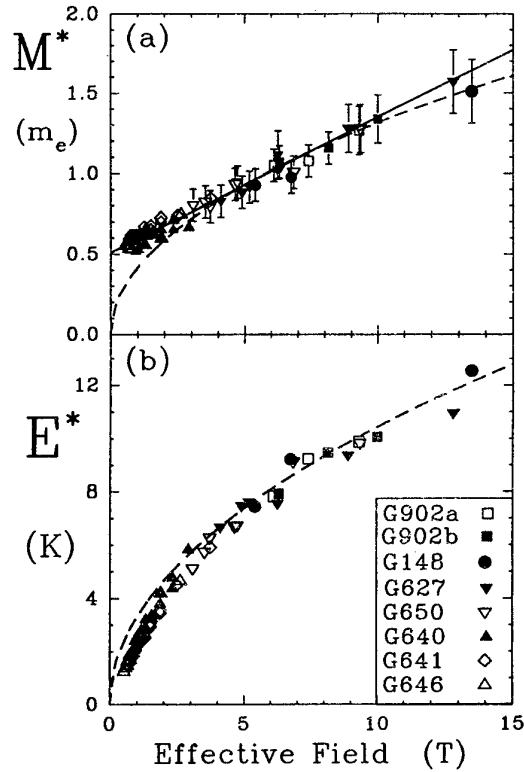


Figure 4. (a) The effective mass as a function of effective field B^* for the family of fractions around $\nu = 1/2$ for a range of different samples with different carrier densities. (b) The CF effective gaps deduced from the effective masses as a function of effective field for the same samples as in (a). The full and broken curves are discussed in the text.

by the use of hydrostatic pressure. Applying pressure decreases the g -factor for electrons via the well known $\mathbf{k} \cdot \mathbf{p}$ perturbation theory expression [15]

$$g^*/2 = \frac{P^2}{3} \left(\frac{1}{E_g} - \frac{1}{E_g + \Delta_0} \right) - 1 \quad (3)$$

where the increase in bandgap due to the application of pressure causes the g -factor to pass through zero in the region of 17 kbar for GaAs. Figure 5 shows the resistivity for a sample with a zero-pressure carrier density of $3 \times 10^{11} \text{ cm}^{-2}$ for four different pressures from 0 to 13.4 kbar. As the pressure increases, the maximum achievable carrier density falls, due to the decrease of the band offset difference between the GaAs and the AlGaAs barrier, and by 13.4 kbar has almost halved. The curves are shown normalized to the level occupancy. Several features are apparent. First, the fractions with odd values of the numerator show a decrease in strength which is particularly pronounced in the region $1 < \nu < 2$. For example the states at 5/3 and 7/5 almost completely disappear by 13.4 kbar. Secondly, there is an enhancement for most of the even-numerator states, and new states appear such as 4/5 and 6/5. Finally there is a decrease in the width of the resistivity minimum at $\nu = 1$. This behaviour illustrates the important role played by the Zeeman energy in determining the behaviour of the different states. By

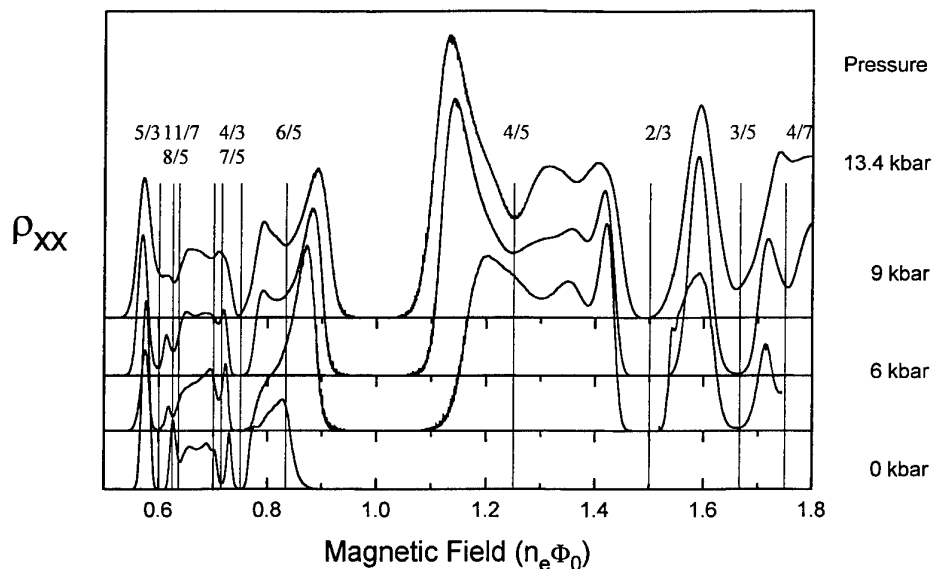


Figure 5. The resistivity as a function of normalized magnetic field for a GaAs/AlGaAs heterojunction with a carrier density of $3.0 \times 10^{11} \text{ cm}^{-2}$ at zero pressure, which falls to 2.1 , 1.87 and $1.6 \times 10^{11} \text{ cm}^{-2}$ at pressures of 6, 9 and 13.4 kbar.

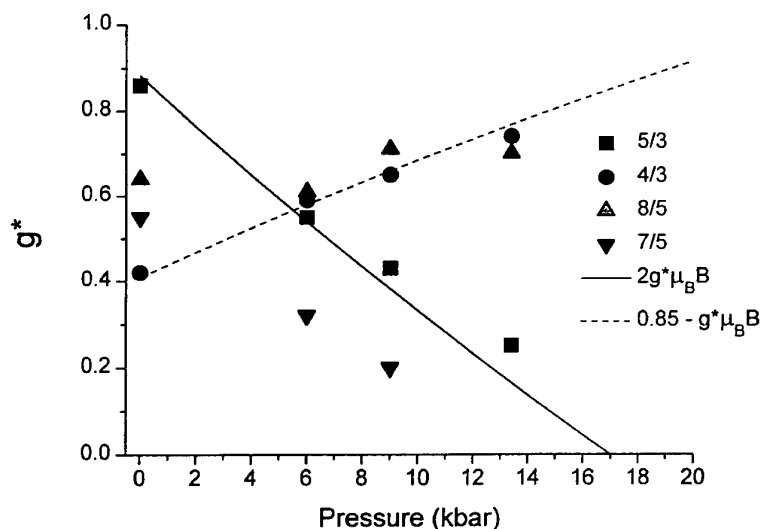


Figure 6. The effective g -factor as a function of pressure, deduced by putting the CF energy gap $E_c^* = g^* \mu_B B$, where B is the total magnetic field.

13.4 kbar we estimate using equation (3) that the GaAs g -factor has fallen to a value of order 0.1 or less, which means that for the sample studied here the Zeeman energy is only of order 0.3 K in the region of $\nu = 3/2$, much less than the activation energies of all of the fractions studied. Under these conditions it appears as if the system is behaving as doubly degenerate 2D layer, with strong fractions only given by two times the usual occupancies.

To analyse the pressure dependence of the energy gaps, we use the same level scheme as proposed by Du *et al* [14], which indicates that 5/3 corresponds to the Zeeman energy gap between the last two CF Landau levels of the two spin states with a polarized ground state, while 4/3

is an unpolarized level with an energy gap determined by the difference of the CF cyclotron energy and the Zeeman energy. Measuring the gaps by fitting the temperature dependence of the oscillation minima as described above allows us to define an effective g -factor for the system from $g^* \mu_B B = E_c^*$, where it should now be remembered that we are using the total magnetic field, and not B^* . The results of such an analysis are shown in figure 6 for several of the minima around $\nu = 3/2$. The effective g -factor measured from 5/3 shows a steady decrease with pressure, following quite closely the calculated decrease in the single-particle value for GaAs, but assuming an exchange enhancement [16] by a factor of two due to the polarization of the system.

By contrast $4/3$ shows a steady increase [17], due to the fall in Zeeman energy, and since this state is unpolarized the increase follows the calculated values using the bare value of the bulk g -factor. Further confirmation of this picture comes from a study of the carrier concentration dependence of the magnitude of the $5/3$ gap [18], which increases almost linearly with carrier density, and thus total magnetic field, consistent with a constant value of the g -factor of 0.85.

Returning to the minimum at $\nu = 1$, we see that there is a steady decrease in width with pressure, but the rate of fall is considerably less than the decrease in the magnitude of the calculated Zeeman energy. This behaviour is probably related to the formation of a residual Coulomb gap even for zero g -factor, as found in double-layer systems [19].

These preliminary measurements indicate that a full understanding of the role of the Zeeman energy in determining the CF properties will be vital.

References

- [1] Tsui D C, Stormer H L and Gossard A C 1982 *Phys. Rev. Lett.* **48** 1559
- [2] Chakraborty T and Pietilainen P 1995 *The Fractional Quantum Hall Effect* 2nd edn (New York: Springer)
- [3] Jain J K 1992 *Adv. Phys.* **41** 105
- [4] Halperin B I, Lee P A and Reed N 1993 *Phys. Rev. B* **47** 7312
- [5] Leadley D R, Nicholas R J, Foxon C T and Harris J J 1994 *Phys. Rev. Lett.* **72** 1906
- [6] Du R R, Stormer H L, Tsui D C, Pfeiffer L N and West K W 1994 *Solid State Commun.* **90** 71
- [7] Manoharan H C, Shayegan M and Klepper S J 1994 *Phys. Rev. Lett.* **73** 3274
- [8] Du R R, Stormer H L, Tsui D C, Yeh A S, Pfeiffer L N and West K W 1994 *Phys. Rev. Lett.* **73** 3274
- [9] Ando T 1974 *J. Phys. Soc. Japan* **37** 1233
- [10] Leadley D R, van der Burgt M, Nicholas R J, Foxon C T and Harris J J 1996 *Phys. Rev. B* **53** 2057
- [11] Haug R J, von Klitzing K, Nicholas R J, Maan J C and Weimann G 1987 *Phys. Rev. B* **36** 4528
- [12] Eisenstein J P, Stormer H L, Pfeiffer L N and West K W 1989 *Phys. Rev. Lett.* **62** 1540
- [13] Clark R G, Haynes S R, Suckling A M, Mallett J R, Wright P A, Harris J J and Foxon C T 1989 *Phys. Rev. Lett.* **62** 1536
- [14] Du R R, Yeh A S, Stormer H L, Tsui D C, Pfeiffer L N and West K W 1995 *Phys. Rev. Lett.* **75** 3926
- [15] Hermann C and Weisbuch C 1977 *Phys. Rev. B* **15** 816, 823
- [16] Nicholas R J, Haug R J, von Klitzing K and Weimann G 1988 *Phys. Rev. B* **37** 1294–302
- [17] Morawicz N G, Barnham K, Briggs A, Foxon C T, Harris J J, Najda S P, Portal J C and Williams M L 1993 *Semicond. Sci. Technol.* **8** 333
- [18] Leadley D R, van der Burgt M, Nicholas R J, Gee P J, Singleton J, Harris J J and Foxon C T 1996 *Surf. Sci.* at press
- [19] Suen Y W *et al* 1992 *Phys. Rev. Lett.* **68** 1379
Eisenstein J P *et al* 1992 *Phys. Rev. Lett.* **68** 1383

Evidence for quasi-classical transport of composite fermions in an inhomogeneous effective magnetic field

J H Smet[†], R Fleischmann[‡], D Weiss[§], R Ketzmerick[‡],
R H Blick[‡], G Lütjering[‡], K von Klitzing[‡], T Geisel[‡] and
G Weimann^{||}

[†] Max-Planck Institut für Festkörperforschung, Heisenbergstraße 1,
D-70569 Stuttgart, Germany

[‡] Institut für Theoretische Physik und SFB Nichtlineare Dynamik, Universität
Frankfurt, D-60054 Frankfurt/Main, Germany

[§] Institut für Experimentelle und Angewandte Physik, Universität Regensburg,
D-93040 Regensburg, Germany

^{||} Fraunhofer Institut für Angewandte Festkörperphysik, D-79108 Freiburg, Germany

Abstract. We report on transverse magnetic focusing experiments in arrays of equidistant constrictions with centre-to-centre constriction spacings L varying from 500 nm to 1.2 μm in the vicinity of the Landau level filling factor $\nu = \frac{1}{2}$ of a two-dimensional electron system. Provided the composite fermion mean free path is comparable to L , focusing is observed for both effective magnetic field directions. The non-uniform electron distribution with a saddle point in the middle of each constriction, and the associated inhomogeneous effective magnetic field for the composite fermions, invokes characteristic differences in the composite fermion focusing spectra compared with the electron focusing spectra near zero magnetic field. These changes are well predicted by quasi-classical calculations of the composite fermion dynamics in the corresponding potential landscape.

1. Introduction

The composite fermion (CF) picture of the fractional quantum Hall effect (FQHE), initiated by Jain [1], attempts to reconcile the apparent phenomenological similarity between the fractional and integer quantum Hall effects (IQHE), despite their entirely different microscopic origin. While the IQHE can be explained satisfactorily in terms of non-interacting electrons in a magnetic field, the FQHE exists only because of electron–electron interactions. The CF model asserts that a quantum Hall liquid of electrons at filling factor $\nu = \frac{1}{2}$ can be described in a mean field sense as a Fermi liquid, with a well defined Fermi surface at zero effective magnetic field, of CFs, each consisting of an electron and two nucleated flux quanta by virtue of the electron–electron interaction [1, 2]. Away from $\nu = \frac{1}{2}$ the CFs experience an effective magnetic field B_{eff} ($B_{\text{eff}} > 0$ for $\nu < \frac{1}{2}$ and $B_{\text{eff}} < 0$ for $\nu > \frac{1}{2}$) and Landau levels of CFs are formed. The series of FQHE liquids may thus be regarded as a manifestation of the IQHE of CFs. Particularly appealing experiments to provide evidence for the existence of CFs are those that also have been used in the past for electrons to show their quasiclassical behaviour in the ballistic transport regime, such as commensurability

oscillations in 1D and 2D periodic structures and transverse magnetic focusing. There is mounting evidence that this CF quasiparticle picture is surprisingly adequate and that quasiclassical dynamics governs their motion [3–6].

We have investigated the magnetoresistance R_{xx} in arrays of constrictions with different spacing. The geometry resembles closely that of a previous experiment by Nakamura *et al* [7] and exhibits large electron transverse magnetic focusing peaks. It lends itself to ensemble averaging in a single measurement to effectively suppress conductance fluctuations which may disturb the observation of CF magnetic focusing peaks [6]. Provided that the constriction spacing remains comparable to the CF mean free path, CF magnetic focusing is observed, confirming that the CF model properly captures the essential physics in the vicinity of $\nu = \frac{1}{2}$. Moreover, to retain optimal resolution of focusing peaks, the constriction width is kept small, causing a saddle-point potential in the constriction region. This translates into an inhomogeneous effective magnetic field for CFs and invokes characteristic changes in the CF focusing spectrum compared to that of electrons. Remarkable agreement is found with quasiclassical simulations of the CF dynamics in such a potential profile.

2. Experimental details

The experiments were performed on a high-mobility GaAs–AlGaAs heterojunction with the 2DEG located 100 nm underneath the sample surface. The samples were shaped into Hall bar geometries using standard lithographic techniques and alloyed AuGe/Ni/Au contact pads. Prior to electron-beam lithography and after brief illumination with a red light-emitting diode (LED), the areal density n_e and electron mobility μ at zero magnetic field and 1.3 K were respectively $1.9 \times 10^{11} \text{ cm}^{-2}$ and $2.4 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. From the longitudinal resistivity at filling factor $\nu = \frac{1}{2}$ one estimates a CF mean free path of 500 nm, contrasting sharply with the electron mean free path of $\approx 17 \mu\text{m}$ at zero magnetic field. Arrays of abutted cavities were defined by electron-beam lithography. As illustrated in figure 1, they consist of three rows in series, each 5 μm apart, of ten square cavities in parallel with lengths L varying from 500 nm to 1.2 μm . After development, the PMMA resist was used as a mask for a reactive ion etch in a SiCl_4 plasma of the ≈ 100 nm wide trenches forming the cavity boundaries. The constrictions of the cavities have a width W of approximately 230 nm. The samples were mounted in a dilution refrigerator with an 18 T superconducting coil and with the external magnetic field perpendicular to the 2DEG. Four-point longitudinal magnetoresistance measurements, as sketched in figure 1(b), of the cavity arrays and an

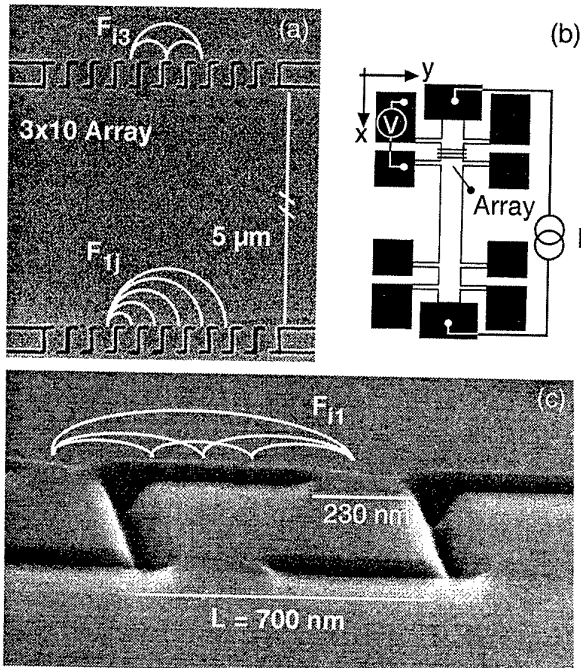


Figure 1. (a) Scanning electron micrograph of a 3×10 square cavity array (only two rows shown) with a period L of 500 nm and a row distance of 5 μm . Representative transverse magnetic focusing trajectories F_{ij} for harmonic index i and subharmonic index j have been added. (b) Measurement configuration for the four-point magnetoresistance $R_{xx} = V/I$. The electrons are focused in a cavity enhance R_{xx} . (c) An enlargement under an angle of 75° of a 700 nm cavity with a lithographic opening width W of 230 nm.

adjacent reference section of the Hall bar were compared at low magnetic fields and in the vicinity of filling factor $\nu = \frac{1}{2}$. During the initial cool-down the constrictions pinched off, but became conducting after illumination with a short red LED light pulse.

3. Experimental results

Figure 2 presents the magnetoresistance R_{xx} of an $L = 500$ nm cavity array in comparison with R_{xx} of a reference section, devoid of cavities, on the same Hall bar. Pronounced differences occur near $B = 0$ and in the vicinity of $\nu = \frac{1}{2}$. The magnetic field position $B_{1/2}$ at half filling was defined from nearby well developed FQHE features with an uncertainty of ± 50 mT. Expanded views near $B = 0$ and around $\nu = \frac{1}{2}$ for the 500 and 700 nm cavity arrays are shown in figure 3. The *electron* transport data were taken for both positive and negative magnetic fields. The resistance maxima around $B = 0$ are clearly identifiable as transverse magnetic focusing. They occur whenever the following matching condition is satisfied between the centre-to-centre constriction spacing L and the electron cyclotron radius R_c^e [8]:

$$2R_c^e i = jL \quad \text{with } i, j = 1, 2, \dots \quad (1)$$

for which electrons ejected from a cavity opening are focused back into another opening, resulting in a resistance maximum. For $j > 1$, electrons skip $j - 1$ openings and are focused into the j th nearest neighbouring cavity.

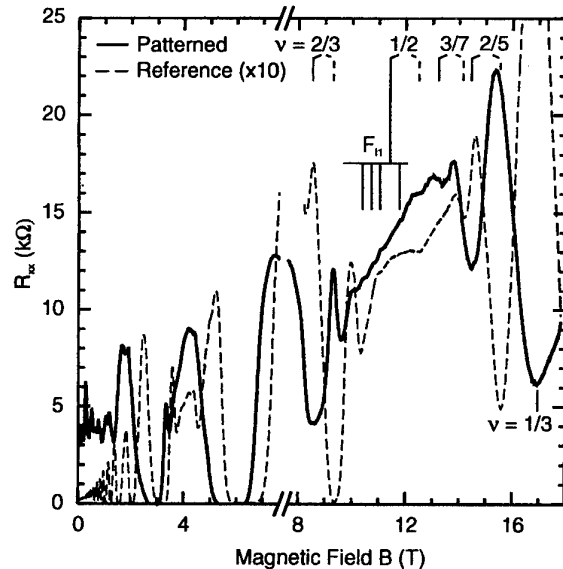


Figure 2. The magnetoresistance R_{xx} of an $L = 500$ nm cavity array (solid curve) is compared with that of the adjacent unpatterned reference section (broken curve) of the Hall bar with a carrier concentration of $1.5 \times 10^{11} \text{ cm}^{-2}$ —after weak illumination—at 180 mK for magnetic fields up to 18 T. The solid trace between 10 and 12 T is enlarged in figure 3(a). The cool-down procedure (for a description see [9]), used to minimize ρ_{xx}^{CF} , accounts for the axis break.

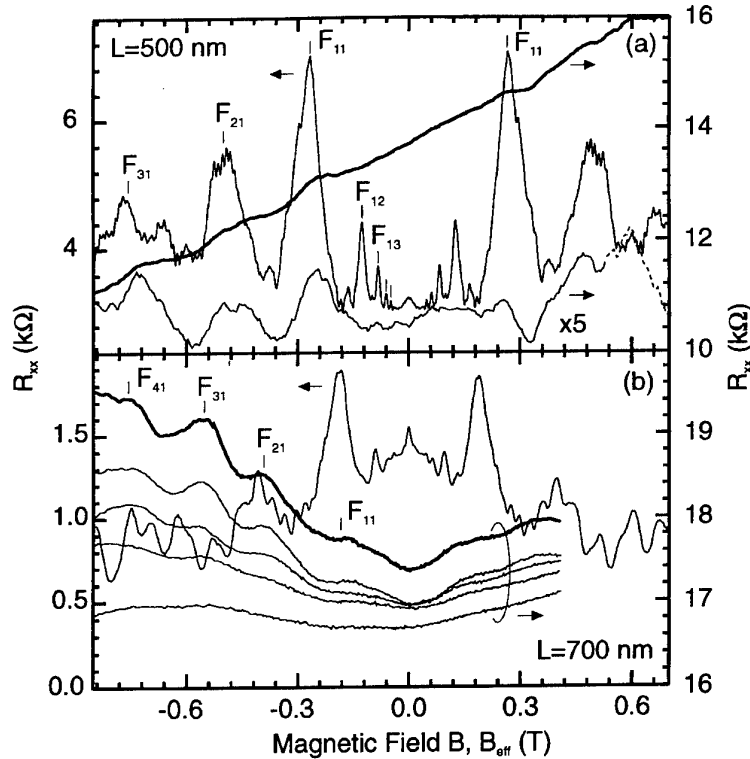


Figure 3. Comparison of the electron (left axes) and CF (right axes) magnetic focusing spectra for $L = 500$ nm and 700 nm cavity arrays with respective carrier concentrations—after different levels of illumination—of 1.4×10^{11} and $2.0 \times 10^{11} \text{ cm}^{-2}$. For the sake of comparison, the magnetic field scale of the CF curves has been divided by $\sqrt{2}$ and shifted horizontally to make $B_{1/2}$ coincide with $B = 0$. The CF traces show pronounced focusing peaks F_{i1} for negative (effective) field ($\nu > 1/2$). For the $L = 1.2 \mu\text{m}$ cavity (not shown) they are absent. Only a weak F_{11} focusing peak can be discerned in (a) for a positive effective field. The bottom curve in (a) has been obtained by subtracting a linear background from the CF curve ($\times 5$, $10 \text{ k}\Omega$ offset). CF curves in (b) do not go beyond $B_{\text{eff}} = 0.5 \text{ T}$ since this corresponds to the maximum field of our magnet. For the CF traces in (b) the less pronounced asymmetry can be ascribed to a reduced saddle potential as expected for a larger n_s value. The temperature dependence of R_{xx} of the $L = 700$ nm cavity array near $\nu = 1/2$ has been indicated (from top to bottom: $100, 320, 580, 700, 830 \text{ mK}$). Subsequent curves have a resistance offset of -200Ω for readability.

Such subharmonics require ballistic transport across larger distances and cannot be resolved for CFs. Therefore, hereafter we will concentrate on the harmonics F_{i1} . For the sake of comparison, the magnetic field axes of the CF traces have been divided by a factor of $\sqrt{2}$, to account for full spin polarization of the CF sea near $\nu = \frac{1}{2}$, and concomitantly shifted downwards to make $B_{1/2}$ coincide with zero magnetic field. One notices that the electron and CF fundamental peak ($i = 1, j = 1$) as well as higher harmonics ($i > 1, j = 1$) line up, confirming the validity of the weakly interacting CF model. For a cavity array with $L = 1.2 \mu\text{m}$ no focusing maxima could be discerned (not shown). The CF focusing features wash out upon raising the temperature above 700 mK , in sharp contrast to the electron resonances which persist up to temperatures as high as 20 K . Despite good agreement of the magnetic field peak positions, inconsistencies between the electron and CF focusing features are obvious to the eye, namely: (i) more harmonics appear near $\nu = \frac{1}{2}$ than around $B = 0$ (for example F_{31} and F_{41} in figure 3(b)); (ii) different strengths of the focusing peaks for positive and negative effective magnetic fields (F_{11} is weaker in figures 3(a) and 2 for $B_{\text{eff}} > 0$); (iii) the weak damping (F_{11} , F_{21} and F_{31} in figure 3(a)) or even increasing instead of decreasing heights

of the focusing peaks for increasingly negative effective field (F_{11} , F_{21} and F_{31} in figure 3(b)).

4. Quasiclassical calculations

The prevailing differences in experiment between CF and electron magnetic focusing originate from the CF property that a non-uniform electrostatic potential $U(x, y)$ and the associated changes in the particle density translate into a spatially modulated effective magnetic field according to [4]

$$B_{\text{eff}}(x, y) = B_{\text{eff}}^0 + B_{1/2} \frac{U(x, y)}{E_F} \quad (2)$$

where E_F is the Fermi energy and $B_{\text{eff}}^0 = B - B_{1/2}$ is the effective magnetic field far away from the *soft* walls in the bulk region. A quasiclassical description of the CF dynamics is feasible because the unknown dispersion of the CFs does not affect their cyclotron radius or the path of their trajectories—even in an arbitrary potential landscape—but only their velocities [4]. Details of the calculations will be given elsewhere [10].

We use a simple three-constriction model with saddle-point potentials in each opening (insets of figures 4 and 5).

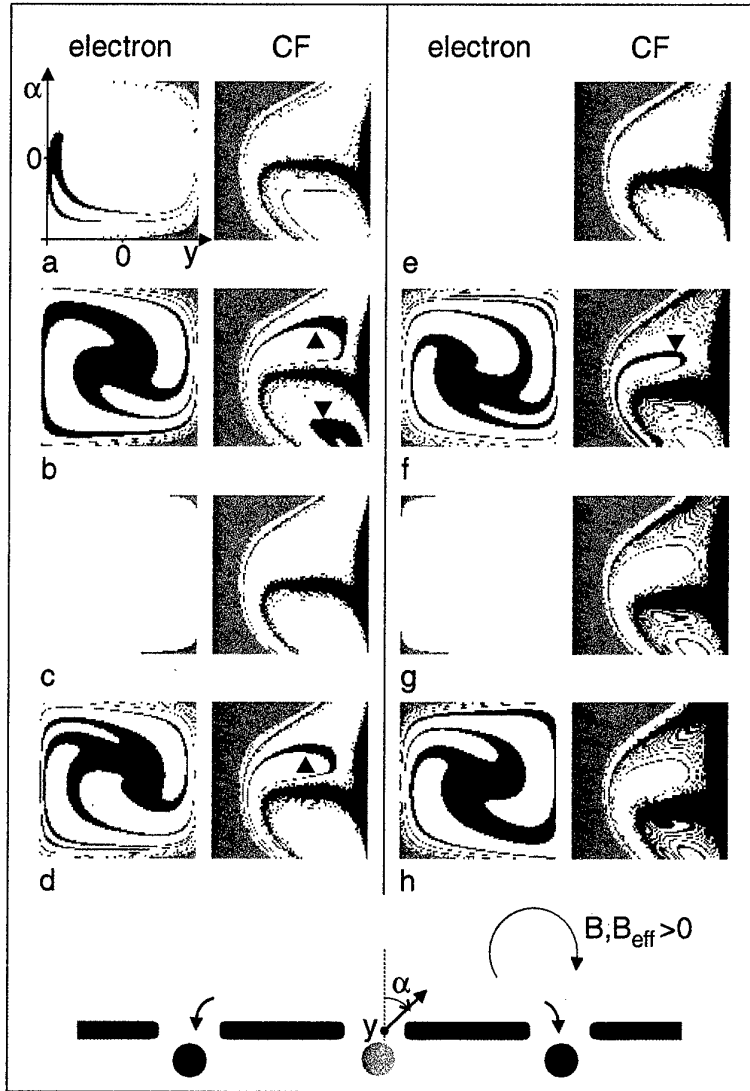


Figure 4. Phase-space portraits in the injection plane of electrons (left) and CFs (right) at eight different magnetic field values (a–h) corresponding to subsequent minima and maxima in the reflection coefficient R as indicated in figure 5. In each square the injection plane is parametrized by the y position (horizontal axes, $-W/2 < y < +W/2$) in the central opening and the injection angle α (vertical axes, $-\pi/2 < \alpha < +\pi/2$). These initial conditions are classified by the final destination of their trajectories and are colour-coded as shown in the inset and explained in the text. Coloured areas associated with the CF-focusing commensurability have been indicated by solid triangles in (b), (d) and (f) (not distinguishable in (h)).

At a fixed external magnetic field, trajectories of particles (either electrons or CFs), injected in the central opening (out of the cavities in experiment) as a function of the injection angle α and as a function of the y position in the central opening, are traced. They can be classified in the following colour-coded bins according to their final destination (figure 4): reflected particles either return to the injector (green) or cross the left (blue) or right (red) constriction; particles that do not return through any of the openings contribute to the net current (white). The coordinate system was chosen such that for an electron (a CF) a positive (positive effective) magnetic field leads to a clockwise rotation of the electron (CF). The bins in injection phase-space of the different trajectory classes constructed according to this colour code are illustrated in

figures 4(a–h) for eight different values of the magnetic field for electrons and the effective magnetic field for CFs. The reflection coefficient R at a specific value of the (effective) magnetic field, i.e. the probability of the trajectories returning through any of the three openings, is obtained by integrating over the coloured areas in the injection plane and appropriate normalization. Figure 5 shows the reflection coefficient for two different saddle-point potentials.

The simulations reveal the following differences for CFs compared with electrons. CFs ejected from the left side of the central opening have a high probability of returning through the same constriction (green phase-space volume, for example for $B_{\text{eff}} = 0$ in figure 4(e)) since the positive effective field related to the density reduction near the wall

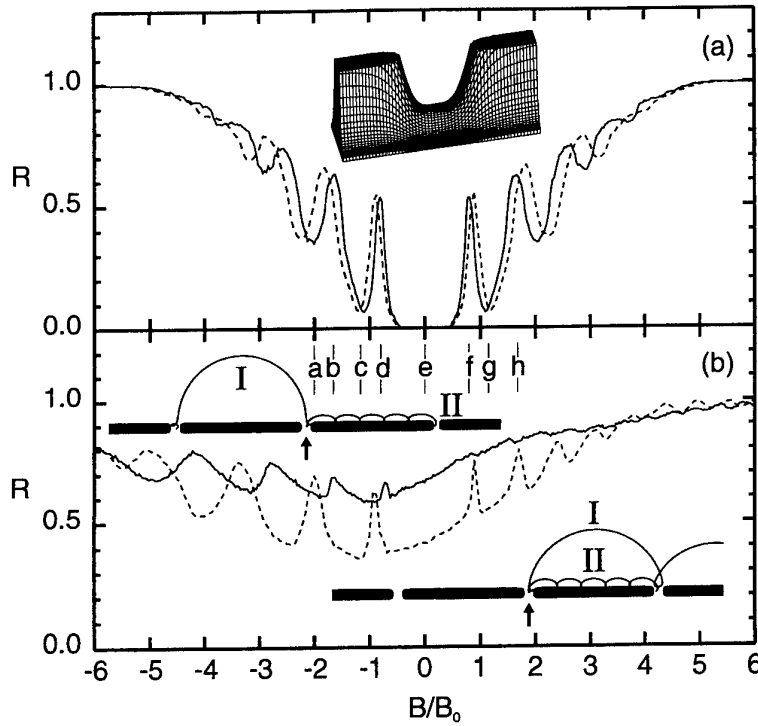


Figure 5. The reflection coefficient R , directly proportional to the coloured areas in figure 4, for particles injected in the central opening marked by an arrow, is shown for (a) electrons versus magnetic field and (b) CFs versus effective magnetic field in units of B_0 corresponding to a cyclotron diameter L far from the constrictions. The three-constriction model potential is described by $U(x, y) = U_0 U_x(x) U_y(y - L) U_y(y) U_y(y + L)$ with $U_x(x) = 1/[1 + (x/\Delta x)^\beta + 1]$, $U_y(y) = 1 - U_1/[1 + (y/\Delta y)^\beta + 1]$, $\beta = 4.0$, $U_0 = 2E_F$, $\Delta x = 0.04L$ and $\Delta y = 0.19L$ (see insets). The widths of the potential wall and of the openings are controlled by Δx and Δy respectively. The saddle-point height is determined by U_1 . Without saddle points in the openings one finds only a very weak asymmetry for CFs between the two field directions as suggested by the calculations for two different saddle-point heights: $U_1 = 0.85$ (solid curve) and $U_1 = 0.925$ (broken curve). The magnetic field values for which the colour-coded injection planes (for the solid curve with $U_1 = 0.85$) are shown in figure 4 have been marked by symbols (a–h). The insets in (b) show the different types of trajectory at $|B_{\text{eff}}^0| \approx 2\hbar k_F/eL$ for positive (right) and negative (left) field directions.

succeeds in bending CF trajectories to their point of origin. Trajectories starting on the right side of the opening are likely to perform a skipping motion along the soft wall and eventually enter the right constriction (red phase-space volume, for example for $B_{\text{eff}} = 0$ in figure 4(e)). These orbits will hereafter be referred to as type II orbits. They are always deflected in the same direction irrespective of the effective field sign. Therefore, the effective opening width is reduced by almost one half, since a large phase-space volume is occupied by orbits that do not contribute to the focusing commensurability. Because a reduction in the effective constriction width enhances the focusing resolution, one anticipates that focusing peaks persist across a larger magnetic field range for CFs than for electrons. This can be seen in figure 5 when comparing the reflection coefficient for electrons and CFs for negative fields and agrees with the experimental observation (i).

For appropriate values of the magnetic field a fraction of particles is focused through a neighbouring constriction and a complex (blue for $B, B_{\text{eff}} < 0$ and red for $B, B_{\text{eff}} > 0$) S-shape develops, signifying that the magnetic focusing condition has been met. The injection plane areas associated with CF focusing are marked by triangles in figures 4(b, d, f). Representative orbits (hereafter referred to as type I orbits) for both magnetic field directions have

been sketched in the insets of figure 5(b). The injection plane area related to CF focusing is *smaller* for $B_{\text{eff}}^0 > 0$ (e.g. compare the marked region in figure 4(f) for $B_{\text{eff}}^0 > 0$ with figures 4(b, d) for $B_{\text{eff}}^0 < 0$). Contrary to type II orbits, type I orbits are deflected in the *opposite* direction when changing the effective field sign. These type I trajectories descend into the flat region and therefore may contribute to the magnetic focusing. However, for *positive* B_{eff}^0 their probability of crossing the saddle-point region into the right-hand opening is small because the positive field is even enhanced in this region. For negative effective fields, B_{eff}^0 and $B_{1/2}$ have opposite signs and the magnitude of $B_{\text{eff}}(x, y)$ at the saddles decreases with decreasing B , allowing more and more CF trajectories of type I to leave the injector and cross the left-hand side opening. These arguments account for the asymmetry and the weak damping or even increasing instead of decreasing height of the focusing peaks for increasingly negative B_{eff}^0 observed in the reflection coefficient R of CFs for the largest saddle-point height (solid curve) in figure 5(b) and the experiment (ii–iii). When the saddle point is lowered, the symmetric behaviour of the CF-focusing peaks is recovered as shown in figure 5(b) (broken curve).

Both electron and CF traces show deviations from the focusing condition in equation (1). The downward shift of

the peaks for electrons is associated with the reduction of the Fermi wavevector k_F^e in the soft-wall and saddle region. The more complex behaviour of the peak position as well as a broadening for CFs reflect the fact that, unlike electrons, CFs are not subject to a constant magnetic field, but rather to a range of magnetic field strengths.

5. Conclusion

In summary, transverse magnetic focusing of CFs in constriction arrays has been observed. The resistance maxima scale properly on the effective magnetic field axis with constriction spacing, and line up with the traditional electron focusing peaks provided a $\sqrt{2}$ scaling factor is included as expected from the full spin polarization near filling factor $\nu = \frac{1}{2}$. Contrary to a previous report [6] but in agreement with conventional expectations, CF focusing could only be discerned when the CF mean free path is large or comparable to the constriction spacing and appears for both negative and positive effective magnetic fields. Finally, quasiclassical calculations of the CF and electron dynamics offer strong support for the idea that the inhomogeneous potential across each constriction and in the vicinity of the cavity walls produces the characteristic changes observed in the experimental CF focusing spectra compared with their electron counterparts.

Acknowledgments

We are indebted to M Riek, F Schartner, B Schönherr and U Waizmann for help with sample preparation. This work has been supported by the Deutsche Forschungsgemeinschaft and the European Community Project 'Human Capital and Mobility', Physics in High Magnetic Fields.

References

- [1] Jain J K 1989 *Phys. Rev. Lett.* **63** 199
- [2] Halperin B I, Lee P A and Read N 1993 *Phys. Rev. B* **47** 7312
- [3] Kang W, Störmer H L and Pfeiffer L N 1993 *Phys. Rev. Lett.* **71** 3850
- [4] Fleischmann R, Geisel T, Holzknicht C and Ketzmerick R 1995 see <http://xxx.lanl.gov/abs/cond-mat/9509168>
- [5] Willet R L, Ruel R R, West K W and Pfeiffer L N 1993 *Phys. Rev. Lett.* **71** 3846
- [6] Goldman V J, Su B and Jain J K 1994 *Phys. Rev. Lett.* **72** 2065
- [7] Nakamura K, Tsui D C, Nihey F, Toyoshima H and Itoh T 1990 *Appl. Phys. Lett.* **56** 385
- [8] van Houten H, Beenakker C W J, Williamson J G, Broekaart M E I, van Loosdrecht P H M, van Wees B J, Mooij J E, Foxon C T and Harris J J 1989 *Phys. Rev. B* **39** 8556
- [9] Kukushkin I V, Haug R J, von Klitzing K and Eberl K 1995 *Phys. Rev. B* **51** R18045
- [10] Fleischmann R, Geisel T and Ketzmerick R 1995 see <http://xxx.lanl.gov/abs/cond-mat/9504063>

Optically pumped nuclear magnetic resonance in the quantum Hall regimes

S E Barrett^{†‡}, G Dabbagh, L N Pfeiffer, K W West and R Tycko^{†§}

AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974, USA

Abstract. We review our recent measurements of the Knight shift ($K_s(\nu, T)$) and spin-lattice relaxation time ($T_1(\nu, T)$) of the ^{71}Ga nuclei located in n-doped GaAs quantum wells using optically pumped NMR, for Landau level filling $0.66 < \nu < 1.76$ and temperature $1.55 \text{ K} < T < 20 \text{ K}$. $K_s(\nu)$ (\propto the electron spin polarization $\langle S_z(\nu) \rangle$) drops precipitously on either side of $\nu = 1$, which is evidence that the charged excitations of the $\nu = 1$ ground state are finite-size skyrmions. For $\nu < 1$, the data are consistent with a many-body ground state which is not fully spin polarized, with a very small spin excitation gap that increases as $\nu \rightarrow 2/3$.

Electron–electron Coulomb interactions are known to significantly affect the physics of two-dimensional electron systems (2DES) in strong magnetic fields, resulting in novel collective phenomena such as the fractional quantum Hall effect [1]. Halperin [2] was the first to point out that these interactions may also lead to significant spin reversal in the many-body ground state and low-lying excited states of 2DES, since for typical experimental conditions (for example $B = 10 \text{ T}$) in GaAs, the Coulomb energy ($E_c = e^2/\kappa L_c = 13.7 \text{ meV}$) is much larger than the Zeeman energy ($E_z = g^*\mu B = 0.3 \text{ meV}$). Subsequent numerical calculations [3, 4] and transport measurements [5–7] support this possibility.

In this paper, we will review our recent optically pumped nuclear magnetic resonance (OPNMR) [8] measurements of the Knight shift K_s [9] and spin-lattice relaxation time T_1 [10] of ^{71}Ga nuclei in an electron-doped multiple quantum well (MQW) structure as a function of Landau level filling factor ν and temperature [11]. OPNMR measurements directly probe the electron spin degree of freedom, providing novel information about the physics of 2DES. For example, the Knight shift measurements provide the first experimental support for the recent predictions that the ground state at $\nu = 1 \pm \epsilon$ contains finite-size skyrmions [12, 13], with an effective spin which reflects the competition between the Coulomb energy and the Zeeman energy. We conclude that OPNMR studies are an important new probe of the many-body ground state and low-lying excited states in 2DES.

The MQW sample used in these measurements contains forty 300 \AA wide GaAs wells separated by 1800 \AA

$\text{Al}_{0.1}\text{Ga}_{0.9}\text{As}$ barriers. Si delta-doping spikes are located in the centre of each barrier. Characterization of this wafer by low-field transport measurements at 4.2 K yielded the electron density $n = (1.41 \pm 0.14) \times 10^{11} \text{ cm}^{-2}$ in each well, with mobility $\mu = 1.44 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. Using home-built NMR probes, the growth axis of the MQW sample was tilted by an angle θ ($0^\circ < \theta < 60^\circ, \pm 0.5^\circ$) away from the constant field B , thereby varying the filling factor ($\nu = nhc/eB \cos \theta$) *in situ*.

The OPNMR measurements described below utilized the timing sequence SAT – τ_L – τ_D – DET, where SAT represents an rf pulse train that saturates (destroys) the nuclear polarization, τ_L is a period of illumination by the laser (σ^+ light, $\lambda = 806 \text{ nm}$, $10\text{--}300 \text{ mW cm}^{-2}$), τ_D is a period of no illumination ($\tau_D \geq 1 \text{ s}$), and DET represents the direct detection of the NMR free induction decay (FID) signals following a single $\pi/2$ pulse [8]. During τ_L , optical pumping of interband transitions generates electrons and holes in the GaAs wells with non-equilibrium spin polarizations, which then polarize the nuclei in the wells through the contact hyperfine coupling. The electronic system quickly equilibrates at the beginning of τ_D , but the enhanced nuclear polarization persists until the DET period, since $T_1 \gg 1 \text{ s}$. In practice, an NMR signal enhancement of about 100 has been obtained using this technique, which makes the study of quantum wells possible.

Figure 1 shows the ^{71}Ga NMR spectra acquired as a function of τ_L ($\tau_D = 1 \text{ s}$) in a 7.05 T field at $T = 1.55 \text{ K}$ with $\theta = 0^\circ$ [9]. The broad, asymmetric resonance observed for $\tau_L = 5 \text{ s}$ is due to the ^{71}Ga nuclei in the GaAs wells, where the non-equilibrium nuclear spin polarization is photogenerated. As τ_L is increased, polarization diffuses into the barriers via the nuclear spin–spin coupling [8, 10]. The narrow, symmetric resonance which dominates the spectrum for $\tau_L = 480 \text{ s}$ is due to the ^{71}Ga nuclei in

[†] To whom correspondence should be addressed.

[‡] Present address: Department of Physics, Yale University, PO Box 208120, New Haven, CT 06520-8120, USA.

[§] Present address: National Institutes of Health, Bldg 5, Rm 112, Bethesda, MD 20892, USA.

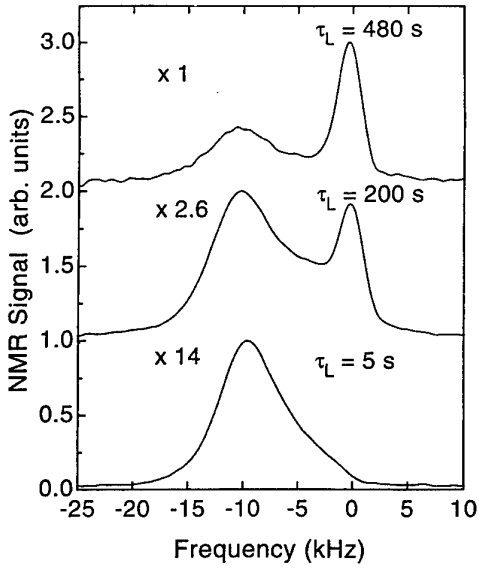


Figure 1. ^{71}Ga NMR spectra of the GaAs/AlGaAs MQW acquired in the dark ($\tau_D = 1$ s) for various optical pumping times τ_L , with $B = 7.05$ T, $\theta = 0^\circ$ and $T = 1.55$ K. For ease of comparison, the signals are scaled by the indicated factors, and are offset for clarity. (Adapted from [9].)

the $\text{Al}_{0.1}\text{Ga}_{0.9}\text{As}$ barriers. The electrons confined in the wells produce an extra hyperfine field which shifts the well resonance below the barrier resonance, as expected for the Fermi contact interaction with $g^* < 0$ [14]. The variation of the lowest subband electron wavefunction along the growth direction leads to the asymmetry of the well resonance [10]. We define the Knight shift to be the peak-to-peak frequency splitting between the well and the barrier resonance lines, so $K_s(\nu, t) = A_{zz} \langle S_z(\nu, T) \rangle$ for $B \parallel z$, where A_{zz} is the hyperfine coupling constant for nuclei in the centre of the well [15].

Figure 2(a) shows the dependence of the ^{71}Ga Knight shift K_s on the tilt angle θ for $B = 7.05$ T and $B = 9.39$ T, at $T = 1.55$ K [9]. Assuming that the maximum K_s observed at $\theta = 28.5^\circ$ for $B = 7.05$ T corresponds to $\nu = 1$, we infer an electron density $n = 1.50 \times 10^{11} \text{ cm}^{-2}$. This density is used to convert the sample tilt angle θ to the Landau level filling factor ν . Figure 2(b) is a plot of K_s versus ν for $B = 7.05$ T and $B = 9.39$ T. The collapse of the two data sets onto one another in figure 2(b) suggests that the hyperfine coupling constant is isotropic, and therefore the Knight shift directly reflects the electron spin polarization, i.e. $K_s(\nu(\theta)) = A \langle S_z(\nu(\theta)) \rangle$. Clearly, $\langle S_z(\nu) \rangle$ drops precipitously both above and below $\nu = 1$.

In figure 3, two fits to the $K_s(\nu)$ data for $B = 7.05$ T at $T = 1.55$ K are shown [9]. The functional form of these fits is obtained by ‘generalizing’ the $T \approx 0$ independent electron model for the spin polarization assuming fixed electron density, which parametrizes the effect of interactions near $\nu = 1$. The underlying assumption is that each quasiparticle (or quasihole) added to the fully polarized $\nu = 1$ ground state results in S (or

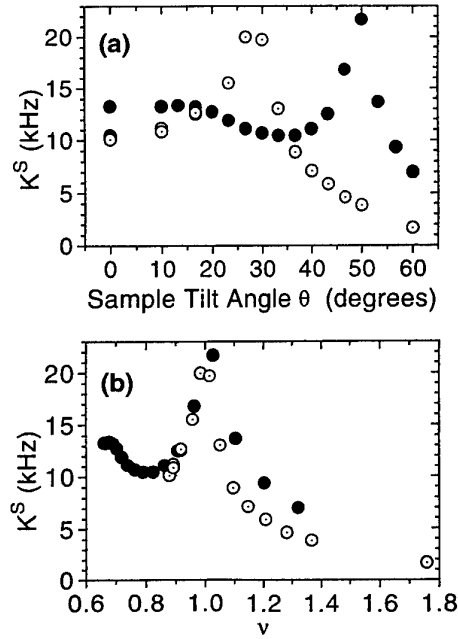


Figure 2. (a) Dependence of ^{71}Ga Knight shift K_s on the sample's tilt angle θ for $B = 7.05$ T (open circles) and $B = 9.39$ T (full circles) at 1.55 K. (b) Dependence of K_s on filling factor ν for $B = 7.05$ T (open circles) and $B = 9.39$ T (full circles) at 1.55 K. The conversion from θ to ν used $n = 1.50 \times 10^{11} \text{ cm}^{-2}$. (Adapted from [9].)

$\mathcal{A} - 1$) flipped spins. In this case

$$K_s(\nu) = \frac{A}{2} \left[\Theta(1 - \nu) \left(\frac{2}{\nu} (1 - \mathcal{A}) - (1 - 2\mathcal{A}) \right) + \Theta(\nu - 1) \left(\frac{2S}{\nu} + (1 - 2S) \right) \right] \quad (1)$$

where $\Theta(x) = 1$ for $x > 0$, and $\Theta(x) = 0$ for $x < 0$. Earlier estimates of the hyperfine coupling imply $A \sim 27$ kHz for three-dimensional electrons [14], or $A \sim 54$ kHz for nuclei in centre of the 300 \AA GaAs well. $A = 44$ kHz is used for the fits in figure 3, a value consistent with the largest shift measured in either field at $T = 1.55$ K. The full curve in figure 3, which assumes $\mathcal{A} = S = 1$ (i.e. non-interacting electrons), fails to fit the data. The broken line is an excellent fit to the data for $0.9 < \nu < 1.1$, with $\mathcal{A} = S = 3.6 \pm 0.3$. Apparently, the charged excitations of the $\nu = 1$ ground state have an effective spin of $(3.6 \pm 0.3)/2$.

These results are in good agreement with the recent predictions that the lowest energy charged excitations of the $\nu = 1$ ground state are finite-size skyrmions, with a calculated effective spin of $7/2$ [13] (or $4.8/2$ [16]) at $E_z/E_c \sim 0.015$. Qualitatively, at $\nu = 1 + \epsilon$, a finite-size skyrmion is cylindrically symmetric with the boundary conditions of a down-spin at $r = 0$ and an up-spin at $r = D$, with a particular radial transition between those two states that maximizes the alignment of nearest-neighbour spins, leaving up-spins for $D < r < \infty$. The length scale D is set by the competition between the Coulomb and Zeeman energy, which increase and decrease D , respectively. The effective spin is $> 1/2$, since all

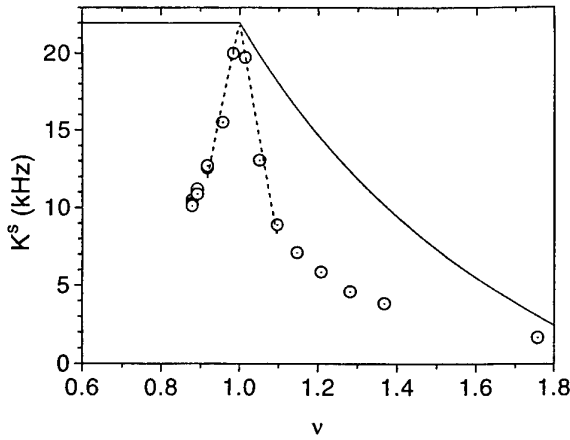


Figure 3. Dependence of K_s on filling factor ν for $B = 7.05$ T (open circles) at 1.55 K. As explained in the text, both fits are given by equation (1), but the full curve has $A = S = 1$ (non-interacting electrons), while the broken line has $A = S = 3.6$ (finite-size skyrmions). (Adapted from [9].)

electrons within $0 < r < D$ are distorted. For brevity, we will refer to this object as a skyrmion [17]. By analogy, the $\nu = 1 - \varepsilon$ ground state is an antiskyrmion, and skyrmion–antiskyrmion pairs are neutral excitations of the $\nu = 1$ ground state. Additional experimental evidence for skyrmions near $\nu = 1$ has recently been reported [18]. The authors measure the electron spin polarization using interband optical transmission. They find that the polarization is dramatically reduced as the filling factor is moved away from $\nu = 1$, in good agreement with the OPNMR Knight shift measurements.

Although skyrmions may occur around other filling factors with incompressible many-body ground states (for example $\nu = 1/3$) [12, 19], they are not expected to be the relevant quasiparticles for $\nu = 3, 5, \dots$ [20]. Recent tilted-field magnetotransport measurements at odd filling factors [21] support these counterintuitive predictions. While there is good agreement between the number of flipped spins per quasiparticle–quasihole pair as determined by transport ($s = 7$) [21] and OPNMR ($s = S + A - 1 = 6.2 \pm 0.6$) [9] at $\nu = 1$ and $E_z/E_c \sim 0.015$, both the magnitude of s and its dependence on E_z/E_c are expected to be somewhat larger based upon Hartree–Fock calculations [13, 16]. Additional theoretical investigations have recently considered a skyrmion crystal [22] and the relationship between skyrmion condensation and certain spin-singlet quantum Hall states [23].

The temperature dependence of the Knight shift probes low-lying excited states, since $K_s(T) \propto \langle S_z(T) \rangle$, and

$$\langle S_z(T) \rangle \equiv \frac{1}{Z} \langle 0 | S_z | 0 \rangle + \sum_i \frac{1}{Z} \exp(-\Delta_i/kT) \langle i | S_z | i \rangle$$

where $|0\rangle$ is the many-body state, Z is the partition function, and the summation is over all excited states $|i\rangle$ with energy Δ_i . Figure 4 (main figure) shows the dependence of K_s on the temperature for $\nu = 0.98$ in the 7.05 T field [9]. The shift saturates at ~ 20 kHz for T below ~ 2 K, which validates our comparison in figure 3 of the $T = 1.55$ K shift data near $\nu = 1$ with the $T \approx 0$ skyrmion model.

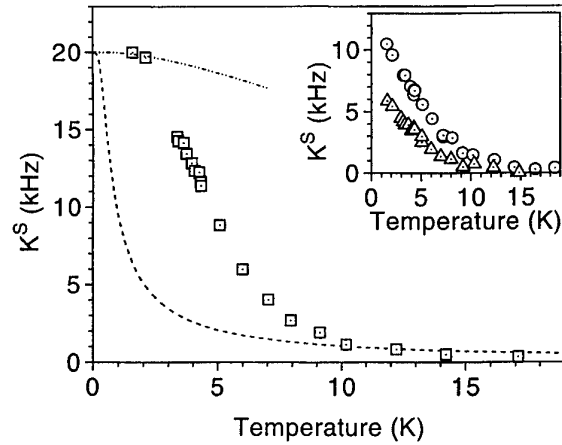


Figure 4. Dependence of K_s on temperature, for $\nu = 0.98$ and $B = 7.05$ T (open squares). As explained in the text, the broken curve is a calculation of $K_s(T)$ assuming non-interacting electrons, and the dash-dot-dot curve is a calculation of $K_s(T)$ assuming that the low-lying excitations are spin waves. Inset: dependence of K_s on temperature, for $\nu = 0.88$ (open circles) and $\nu = 1.2$ (open triangles), at $B = 7.05$ T. (Adapted from [9].)

Figure 4 also contains the prediction for $\langle S_z(T, \nu = 1) \rangle$ assuming: (i) non-interacting electrons (broken curve), or (ii) an interacting 2DES (dash-dot-dot curve) with spin-wave modes as the low-lying excitations [9, 24, 25]. The $\langle S_z(T, \nu = 1) \rangle$ data do not follow either curve. For T above ~ 2 K, $K_s(T)$ drops off more rapidly than the spin-wave fit, which qualitatively suggests a collapse of the exchange energy as the polarization is reduced. Recently, sophisticated calculations that either stress the similarity to an itinerant electron ferromagnet [26] or use a continuum quantum field theory of a ferromagnet [27], have led to predictions for $\langle S_z(T, \nu = 1) \rangle$ which are closer to the data, but significant discrepancies remain. $K_s(T)$ may also be measured at arbitrary ν using OPNMR. The inset of figure 4 shows the data for $\nu = 0.88$ and $\nu = 1.2$, which do not saturate by $T \sim 2$ K, demonstrating the dependence of the many-body states and energy spectrum on the filling factor.

OPNMR also enables measurement of the nuclear spin-lattice relaxation rate $1/T_1$, which probes electron spin dynamics [28, 29]. Figure 5 shows the dependence of $1/T_1$ on filling factor ν at $T = 2.1$ K for $B = 7.05$ T and $B = 9.39$ T [10]. The rate $1/T_1$ is determined by inverting the well resonance for $\tau_L = 5$ s, and monitoring its recovery during τ_D . The $1/T_1$ data are strongly dependent upon ν , but nearly independent of total field, which suggests that the Fermi contact interaction dominates. Since the relaxation involves electron–nucleus flip–flop transitions between nearly degenerate initial and final electron states, a slow (fast) $1/T_1$ implies a large (small) gap to spin-flip excitations [15, 24]. Although the slow rate at $\nu = 1$ is consistent with either interacting or independent electrons, the rapid $1/T_1$ at $\nu \sim 0.85$ can only be explained if interactions induce nearly gapless low-lying spin-flip excitations. The broken line in figure 5 shows that $1/T_1(\nu) \propto (K_s(\nu = 1) - K_s(\nu))$ for $0.9 < \nu < 1.1$, where $K_s(\nu)$ is the broken line in figure 3. This suggests that

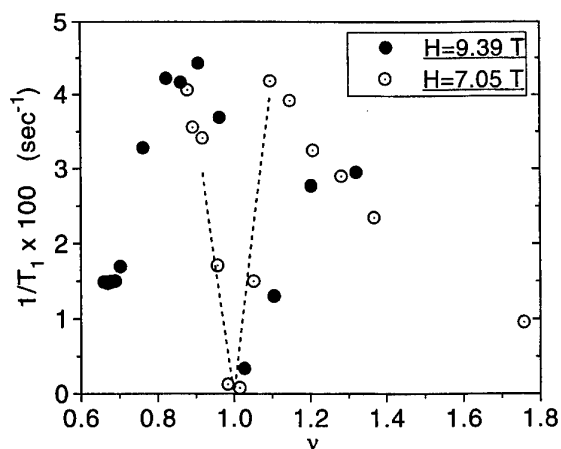


Figure 5. Dependence of ^{71}Ga nuclear spin-lattice relaxation rate $1/T_1$ on filling factor ν at $T = 2.1$ K, for $B = 7.05$ T (open circles) and $B = 9.39$ T (full circles). The broken line is directly proportional to $(K_s(\nu = 1) - K_s(\nu))$, where $K_s(\nu)$ is given by equation (1), with $A = S = 3.6$ (see figure 3). (Adapted from [10].)

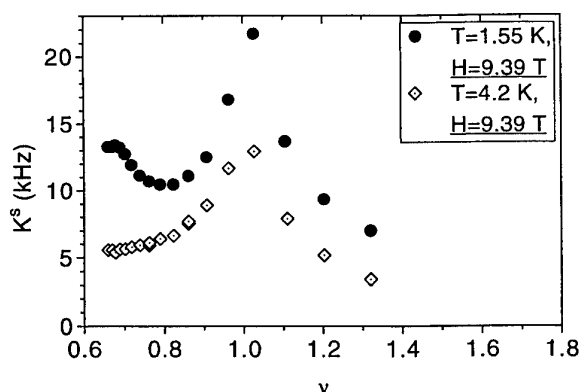


Figure 6. Dependence of K_s on filling factor ν for $B = 9.39$ T, at $T = 4.2$ K (open diamonds) and $T = 1.55$ K (full circles). (Adapted from [9].)

low-lying spin-flip excitations near $\nu = 1$ are related to the number of skyrmions in the ground state.

Clear signatures of the fractional quantum Hall regime [3–5] are also evident in the OPNMR data. First, figure 5 shows that $1/T_1(2.1 \text{ K}, \nu)$ has a local minimum near $\nu = 2/3$, one of the fundamental FQHE states. Second, $K_s(1.55 \text{ K}, \nu)$ has a local maximum near $\nu = 2/3$, as is seen in figure 6 [9]. The fact that $K_s(1.55 \text{ K}, \nu = 2/3) \ll K_s(1.55 \text{ K}, \nu = 1)$ suggests that the $\nu = 2/3$ FQHE ground state may not be fully polarized [3,5] for $B = 9.39$ T, but these measurements need to be repeated at lower temperatures to ensure that $\langle S_z(T, \nu = 2/3) \rangle \simeq \langle S_z(0, \nu = 2/3) \rangle$. Qualitatively, the temperature dependence of $K_s(\nu)$ shown in figure 6 suggests that the energy gap to spin-flip excitations is large, medium and small for $\nu = 1, 2/3$ and ~ 0.9 respectively. This assignment is consistent with our measurements for ^{71}Ga T_1 , since its temperature dependence is a strong function of ν , as is seen in table 1 [9].

Table 1. ^{71}Ga nuclear spin-lattice relaxation time T_1 as a function of temperature and Landau level filling factor ν . (Adapted from [9].)

	$\nu = 1.01$	$\nu = 0.88$	$\nu = 0.66$
$T_1(T = 4.2 \text{ K})$ (s)	122	24	43
$T_1(T = 2.1 \text{ K})$ (s)	1280	20	68

We have shown that OPNMR is a powerful local probe of 2DES. Many interesting features of the above results remain to be understood theoretically. In future work, we will extend these measurements to additional filling factors, lower temperatures, and higher fields.

Acknowledgments

We thank J P Eisenstein, S M Girvin, M Kasner, A H MacDonald, N Read, S Sachdev and S L Sondhi for helpful discussions, and B I Greene and T D Harris for experimental assistance.

References

- [1] Tsui D C, Störmer H L and Gossard A C 1982 *Phys. Rev. Lett.* **48** 1559
- [2] Halperin B I 1983 *Helv. Phys. Acta* **56** 75
- [3] Chakraborty T and Pietiläinen P 1988 *The Fractional Quantum Hall Effect* (Berlin: Springer)
- [4] Prange R E and Girvin S M (ed) 1990 *The Quantum Hall Effect* 2nd edn (New York: Springer)
- [5] Eisenstein J P and Störmer H L 1990 *Science* **248** 1510
- [6] Willett R L, Eisenstein J P, Störmer H L, Tsui D C, Gossard A C and English J H 1987 *Phys. Rev. Lett.* **59** 1776
- [7] Clark R G, Haynes S R, Suckling A M, Mallett J R, Wright P A, Harris J J and Foxon C T 1989 *Phys. Rev. Lett.* **62** 1536
- [8] Eisenstein J P, Störmer H L, Pfeiffer L N and West K W 1989 *Phys. Rev. Lett.* **62** 1540
- [9] Furneaux J E, Syphers D A and Swanson A G 1989 *Phys. Rev. Lett.* **63** 1098
- [10] Barrett S E, Tycko R, Pfeiffer L N and West K W 1994 *Phys. Rev. Lett.* **72** 1368
- [11] Barrett S E, Dabbagh G, Pfeiffer L N, West K W and Tycko R 1995 *Phys. Rev. Lett.* **74** 5112
- [12] Tycko R, Barrett S E, Dabbagh G, Pfeiffer L N and West K W 1995 *Science* **268** 1460
- [13] For another review of these experiments see Barrett S E, Dabbagh G, Pfeiffer L N, West K W and Tycko R 1996 *Surf. Sci.* **361/362** 261
- [14] Sondhi S L, Karlhede A, Kivelson S A and Rezayi E H 1993 *Phys. Rev. B* **47** 16 419
- [15] Fertig H A, Brey L, Côté R and MacDonald A H 1994 *Phys. Rev.* **50** 11 018
- [16] Paget D, Lampel G, Sapoval B and Safarov V I 1977 *Phys. Rev. B* **15** 5780
- [17] Slichter C P 1990 *Principles of Magnetic Resonance* 3rd edn (New York: Springer)
- [18] Karlhede A and Sondhi S L unpublished
- [19] Lee D-H and Kane C L 1990 *Phys. Rev. Lett.* **64** 1313
- [20] Rajaraman R 1989 *Solitons and Instantons* (Amsterdam: North-Holland)
- [21] Ladbury R 1995 *Physics Today* July 19
- [22] Aifer E H, Goldberg B B and Broido D A 1996 *Phys. Rev. Lett.* **76** 680
- [23] MacDonald A H private communication

- [20] Jain J K and Wu X-G 1994 *Phys. Rev. B* **49** 5085
Wu X-G and Sondhi S L 1995 *Phys. Rev. B* **51** 14 725
- [21] Schmeller A, Eisenstein J P, Pfeiffer L N and West K W
1995 *Phys. Rev. Lett.* **75** 4290
- [22] Brey L, Fertig H A, Côté R and MacDonald A H 1995
Phys. Rev. Lett. **75** 2562
- [23] Nayak C and Wilczek F 1995 *Preprints LANL*
cond-mat/9501052, cond-mat/9505081, ✓
cond-mat/9507016
- [24] Narath A 1967 *Hyperfine Interactions* ed A J Freeman and
R B Frankel (New York: Academic) ch 7
- [25] Bychkov Y A, Iordanskii S V and Eliashberg G M 1981
JETP Lett. **33** 143
Kallin C and Halperin B I 1984 *Phys. Rev. B* **30** 5655
- [26] Kasner M and MacDonald A H 1995 *Physica B* **212** 289;
1996 *Phys. Rev. Lett.* **76** 3204
- [27] Read N and Sachdev S 1995 *Phys. Rev. Lett.* **75** 3509
- [28] Berg A, Dobers M, Gerhardt R R and von Klitzing K
1990 *Phys. Rev. Lett.* **64** 2563
- [29] Vagner I D and Maniv T 1988 *Phys. Rev. Lett.* **61** 1400

Single-electron transistors with a self-assembled quantum dot

M Dilger, R J Haug, K Eberl and K von Klitzing

Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1, 70569 Stuttgart, Germany

Abstract. Single-electron transistors with a self-assembled quantum dot are fabricated by direct epitaxial growth on patterned substrates. A self-assembled quantum dot is formed during the deposition of an $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ heterostructure on a bow-tie-shaped constriction, pre-patterned on a GaAs substrate. The self-assembled quantum dot is located in the centre of the constriction and is coupled via tunnelling barriers, also fabricated within the same growth process, to electrical leads. The fabricated devices allow switching with single electrons up to temperatures of 6 K by applying a voltage to an in-plane gate, which is also fabricated during the epitaxial growth. The structure of the devices is also characterized using scanning-electron and atomic-force microscopy.

1. Introduction

Quantum dots are three-dimensionally confined electronic systems, whose discrete energy spectra permit storage of single electrons [1, 2]. In single-electron transistors (SETs), quantum dots are coupled via tunnelling barriers to electrical leads, so that switching with single electrons is possible. In most of the SETs prepared in semiconducting material up to now, e.g. [3], the field effect is used to define the quantum dot and the tunnelling barriers. Typical diameters of such quantum dots are about 100 nm. Epitaxy with self-organization has been utilized to grow quantum dots with diameters of merely a few nanometres [4–8]. However, until now only optical and capacitance measurements of arrays of such directly grown quantum dots have been carried out [8–10]. Here we present a novel self-organization scheme to realize a single quantum dot including self-aligned tunnelling junctions to electrical leads, by direct epitaxial growth on patterned substrates. Our method allows electrical measurements of single structures formed at a defined position, in contrast to the large number of quantum dots involved in previous work. One growth process is sufficient to create a very compact SET tunable with a voltage on a single gate, which is also fabricated within the epitaxial growth.

2. Sample fabrication

For the self-organized formation of a quantum dot, we use the surface selectivity of molecular-beam epitaxial (MBE) growth on patterned substrates [11–14]. The pre-patterning of a (100) GaAs substrate, which triggers the formation of the quantum dot at a defined position, consists of two grooves, 1.5 μm wide and 2 μm deep, with sharp rectangular edges, fabricated by reactive-ion etching

(figure 1). The two grooves define a constriction with the shape of a bow-tie and a minimal width of 1 μm . The four groove sections forming the constriction are oriented along $\langle 001 \rangle$ directions to provide isotropic growth conditions. The etch mask is defined using electron-beam lithography, but may also be realized by optical lithography. Subsequently, in the MBE chamber, an $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ heterostructure is deposited at a substrate temperature of 580 °C. The heterostructure has a total thickness of 284.5 nm, consisting of a 10 nm GaAs buffer layer, 35 periods of a 1.9 nm AlAs/3.8 nm GaAs superlattice, a 20 nm GaAs quantum well, a 10 nm undoped $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ spacer, a 40 nm Si-doped $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ layer, covered by a 5 nm GaAs cap. The epitaxial layers on different mesas are not connected, because of the 2 μm deep etched grooves. Therefore gate operation without leakage current is possible between layers on different sides of the grooves (see figure 1(c)). Low-temperature measurements at 4.2 K show that the gate-drain current, I_{GD} , is below 10 pA up to a gate voltage V_G of 10 V.

A quantum dot, buried under a pyramidal structure, is formed on the narrowest part of the ridge between the two grooves, due to surface diffusion processes of deposited atoms during growth. On the wider regions of the ridge a conventional two-dimensional electron gas (2DEG) is formed. The 2DEG on the ridge is coupled to the quantum dot via tunnelling barriers, formed in the transition area between the wide and the narrow parts. After the epitaxial growth, a transistor structure (see figure 1(a)) is etched out of the heterostructure using optical lithography, so that drain, source and gates can be contacted separately. The contacts to the 2DEG are also achieved by means of optical lithography (figure 1(a)).

The pyramid with the buried quantum dot (figures 1(b) and 1(d)) in the centre of the pre-patterned bow-tie-shaped constriction is created by differences in growth rates of the

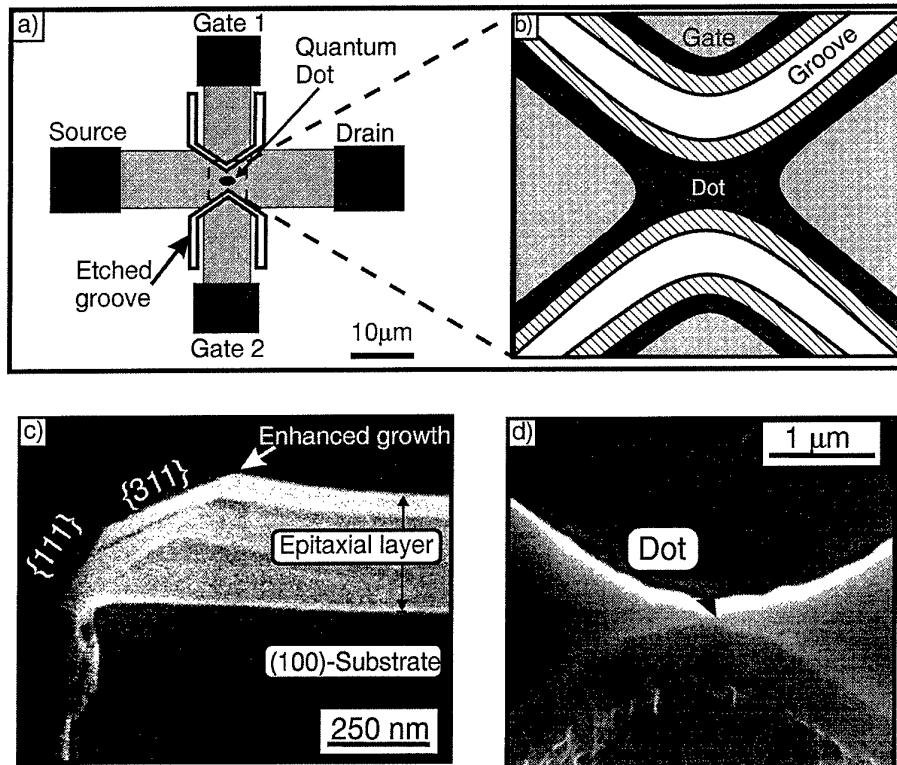


Figure 1. Schematic illustration (a, b) and characterization by SEM (c, d) of the fabricated SETs. (b, c) Stripes with a larger layer thickness (dark regions in (b)) evolve during the growth parallel to the grooves. The basis for the formation of the self-assembled quantum dot is the overlap of two stripes of enhanced growth in the centre of the constriction. A pyramidal structure is caused thereby, under which a quantum dot is buried (d).

epitaxial layer. The surface energy of the epitaxial layer is minimized through facets which are tilted with respect to the (100) surface of the substrate (figure 1(c)). These facets evolve during epitaxial growth along the edges of the patterned substrate. Which of these facets evolve depends on the orientation of the substrate patterning [16]. Surface-diffusion processes of Ga atoms reduce the growth rate of GaAs on the tilted facets and enhance it on the neighbouring regions on (100) facets [15, 16]. This effect is weaker for AlAs due to the much smaller surface diffusion of Al atoms. The cross-sectional scanning-electron microscopy (SEM) photograph of an overgrown sample (figure 1(c)) reveals the thickness modulation of the epitaxial layer near an edge of the patterned substrate. We determined an additional layer thickness of 35 nm due to the enhanced growth near a single edge from the cross-sectional scanning-electron micrograph (see figure 1(c)). The highest growth rate of GaAs in the [100] direction is achieved at the narrowest part of the ridge, caused by contributions from two neighbouring tilted facets, creating the pyramidal structure. The structure in the centre of the constriction with the buried quantum dot is shown in the SEM micrograph of figure 1(d).

The formation of the buried quantum dot and the tunnelling barriers is explained by means of two perpendicular cross-sections through the centre of the pyramid (figure 2). In the narrowest part of the ridge, in the centre of the constriction, the growth rate of GaAs on the [100] plateau is strongly enhanced by surface diffusion

from adjacent {311} facets. The enhanced growth of GaAs in the centre of the constriction shown in figure 2(a) also causes a modulation of the layer thicknesses in the direction from drain to source, shown in figure 2(b). Thereby, an elevated plateau is formed in the middle of the constriction during the growth of a superlattice, surrounded by inclines not only in the direction of the grooves (figure 2(a)) but also in the perpendicular direction (figure 2(b)). Diffusion of Ga atoms from these inclines to the (100) plateau during further growth causes an additional enhancement of the growth rate of GaAs there. If a GaAs quantum well is now deposited, a 'fourfold' enhanced growth occurs on top of the plateau, due to diffusion from neighbouring inclines in all four directions, forming a quantum dot. On the other hand, reduced growth occurs on the inclines in the direction of drain and source and thus generates tunnelling barriers between the quantum dot and the leads (figure 2(b)). Summarizing these points, the self-organization phenomenon consists of three main steps: (1) the formation of additional inclines, (2) reduced growth on these inclines, generating tunnelling barriers, and (3) a 'fourfold' enhanced growth on the plateau on top of the inclines, forming a quantum dot. Subsequent deposition of a modulation Si-doped $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ layer and a GaAs cap embeds the active structure. The silicon doping of the $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ layer causes an n-type doping of the GaAs quantum well on the (100) planes of the structure. It is at present not possible to estimate if the doping of the

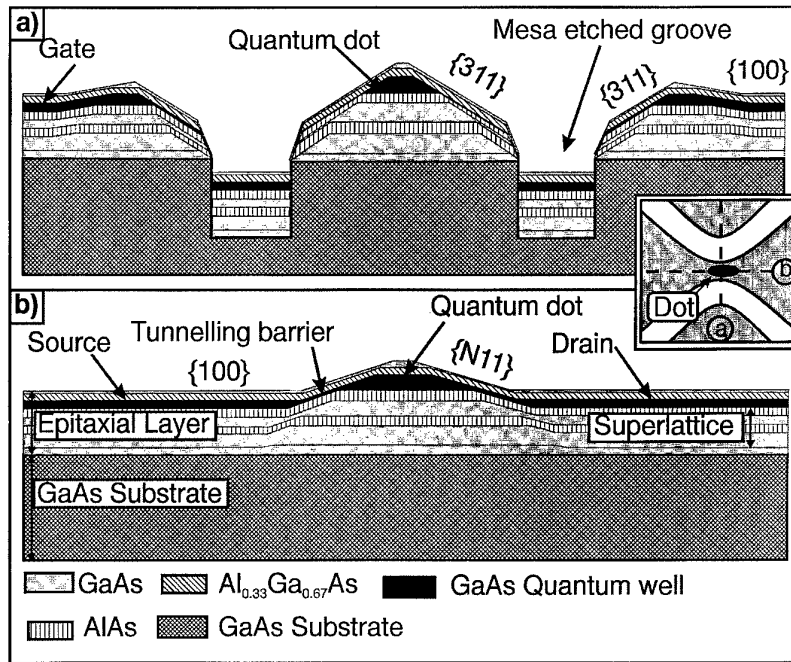


Figure 2. Schematic cross-sections through an SET along the two perpendicular dashed lines shown in the inset (see also figure 1). The modulation of the GaAs quantum well provides the main building blocks of an SET in a self-aligned way. Following the quantum well in (b) from the left to the right side, a conventional 2DEG, used as a source contact, is first of all formed. The following thinner part of the layer on the tilted $\{N11\}$ facet acts as a tunnelling barrier. The thick elevated GaAs island represents the quantum dot, followed again by a tunnelling barrier, directly connected to a 2DEG used as a drain contact.

thin GaAs quantum well on the tilted $\{311\}$ facets on the sidewalls to the grooves is n-type or p-type.

The enhanced growth rate at the centre of the constriction due to the self-organization process is confirmed by atomic-force microscopy (AFM) (figure 3). The method allows the measurement of the height profile in addition to the normal in-plane resolution of a top view. A surface scan along the line drawn in the top view of figure 3(a), i.e. in the drain–source direction, is shown in figure 3(b). It shows that the layer in the centre of the constriction is about 100 nm thicker than the nominal layer thickness, i.e. more than twice the additional layer thickness along a single edge (figure 1(c)). The surface scan demonstrates the thickness modulation of the epitaxial layer due to surface diffusion of deposited atoms and confirms the growth model explained using figure 2.

3. Transport investigations

Electronic-transport measurements (figure 4) reveal that the self-organized quantum dot can be charged and discharged with single electrons. The conductivity through the quantum dot versus the voltage of the tuning gates shows pronounced periodic oscillations up to a temperature of 5.9 K (inset of figure 4(a)). Increasing the gate voltage from one conductance minimum to the next increases the electron number on the quantum dot only by one, due to the small capacitance and therefore large charging energy of the small structure. Figures 4(a) and 4(b) show separate measurements on this device in a $^3\text{He}/^4\text{He}$

dilution refrigerator at a base temperature of 22 mK, which exhibit clearly resolved conductivity peaks, separated by regions of zero conductivity ('Coulomb blockade'). The voltage change required to add an electron to the quantum dot is the same for the two different measurements, one done at 1.3 K (inset of figure 4(a)) and one done at 22 mK after warming up the sample to room temperature (figure 4(a)). We conclude that the quantum dot is geometrically defined during the growth process, since the capacitance between the quantum dot and the gates remains the same, even if the sample is heated up between two measurements. The Coulomb energy is determined to be $E_C = 2.5 k_B T = 1.27$ meV, from the temperature at which the Coulomb-blockade oscillations vanish. The measurements in figures 4(a) and 4(b) show that the period of the Coulomb-blockade oscillations doubles if the gate voltage V_G is applied to only one side gate (figure 4(a)) instead of both side gates (figure 4(b)). This doubling of the Coulomb-blockade oscillations shows that the self-assembled quantum dot is located in the centre of the constriction. This shows furthermore that both side gates allow a similar tuning of the quantum dot. The non-uniform period of the Coulomb-blockade oscillations may be caused by the complex shape of the direct-grown quantum dot, but is up to now not clearly understood. Spectroscopic transport measurements, including measurements with finite drain–source voltage V_{DS} (see e.g. inset of figure 4(b)), are used to determine the regions of Coulomb blockade in the V_{DS} versus V_G plane, which allows the calculation of the Coulomb energy [17]. The Coulomb energy is found to be $E_C = 1.21$ meV. The diameter of the quantum dot is thus

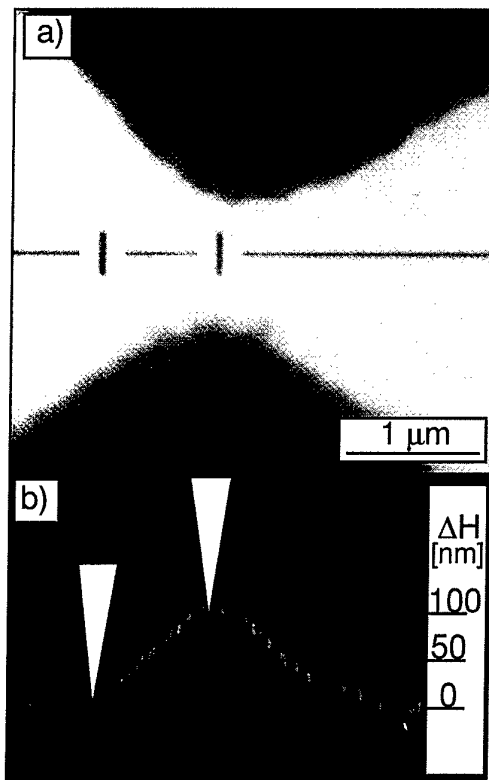


Figure 3. (a) AFM image of a sample after MBE growth. (b) Line scanning along the line shown in (a). The elevation in the centre of the constriction is approximately 100 nm. The two markers correspond to the two vertical strokes in the top view of (a). They are used as a guide to the eye to make the abrupt ascent of the surface in the direction of the centre of the constriction clear.

estimated to be about 100 nm, assuming that the quantum dot is a sphere.

4. Conclusion and outlook

In conclusion, we present a novel self-organization phenomenon for the fabrication of a complete SET by one MBE growth step. The shape of the direct-grown single-electron structure depends on the different growth rates in the centre of the constriction and the neighbouring regions. Therefore the shape of the device can be controlled via the different growth rates which are caused by surface diffusion processes of deposited atoms. The differences in the growth rates can be controlled by varying the material composition, the growth temperature or the shape of the pre-patterning as this also influences the surface diffusion processes. In particular the angle between the groove sections defining the bow-tie-shaped constriction allows us to control the length of the tunnelling barriers independently of the size of the quantum dot. By applying our method to other materials and strained heterostructures including InP or $\text{In}_2\text{Ga}_{1-z}\text{As}$, which show a higher surface diffusion than GaAs, it will be possible to realize even smaller quantum dots. The size of those quantum dots is controlled within the epitaxial growth and is not restricted to the resolving power of the

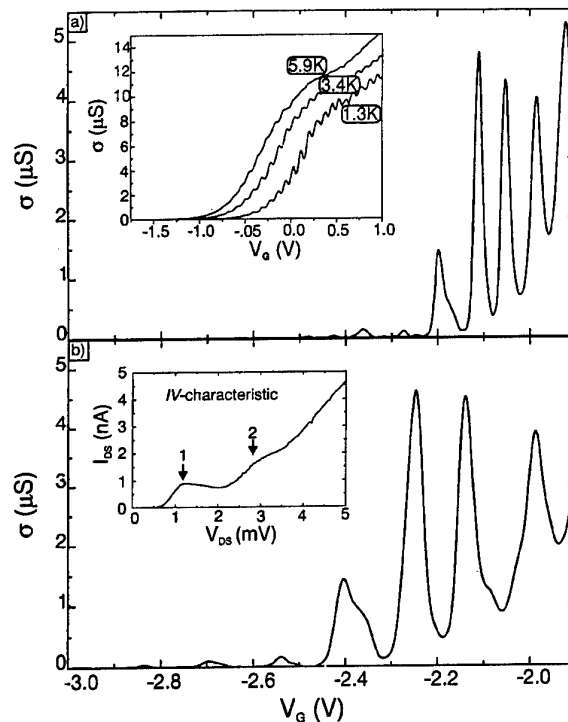


Figure 4. The conductance σ through the quantum dot versus the voltage V_G at the tuning gates. For the measurements shown in (a) the voltage V_G is applied to both side gates of the device, whereas for the measurements shown in (b) the voltage at one gate is fixed to -2.0 V and the voltage V_G is applied only to one side gate, which leads to a doubling of the period of the Coulomb-blockade oscillations. Inset in (a): measurements at higher temperatures, which show that even at temperatures of 6 K a modulation of the σ - V curve is visible. Inset in (b): the current-voltage staircase of the fabricated device, measured at a base temperature of 22 mK.

lithography. Therefore conventional optical lithography can be used for device fabrication. The fabricated device shows that it is possible to realize SETs which operate like normal transistors by tuning only one gate.

Acknowledgments

We would like to thank A Kurtenbach for his contributions to the device fabrication, Y Kershaw for the AFM investigations, J Weis for helpful discussions, and H J Queisser and R Martin for a critical reading of the manuscript. The expert technological help of B Schönherr, F Scharfner and C Lange is gratefully acknowledged. This work was supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.

References

- [1] 1991 Special issue on single charge tunnelling *Z. Phys. B* **85** (3)
- [2] Grabert H and Devoret M H (eds) 1992 *Single Charge Tunneling* (NATO ASI Series B 294) (New York: Plenum)

- [3] Weis J, Haug R J, von Klitzing K and Ploog K 1993 *Phys. Rev. Lett.* **71** 4019
- [4] Merz J L and Petroff P M 1991 *Mater. Sci. Eng.* **9** 275
- [5] Leonard D, Krishnamurthy M, Reaves C M, Denbaars S P and Petroff P M 1993 *Appl. Phys. Lett.* **63** 3203
- [6] Moison J M, Houzay F, Barthe F, Leprince L, Andre E and Vatel O 1994 *Appl. Phys. Lett.* **64** 196
- [7] Gossard A C and Fafard S 1994 *Solid State Commun.* **92** 63
- [8] Nötzel R, Temmyo J and Tamamura T 1994 *Nature* **369** 131
- [9] Drexler H, Leonard D, Hansen W, Kotthaus J P and Petroff P M 1994 *Phys. Rev. Lett.* **73** 2252
- [10] Kurtenbach A, Eberl K and Shitara T 1995 *Appl. Phys. Lett.* **66** 361
- [11] Kapon E, Tamargo M C and Hwang D H 1987 *Appl. Phys. Lett.* **54** 347
- [12] Fukui T and Ando S 1989 *Electron. Lett.* **25** 410
- [13] Nilsson S, Van Gieson E, Arent D J, Meier H P and Forster T 1989 *Appl. Phys. Lett.* **55** 972
- [14] Nakamura Y, Koshiha S, Tsuchiya M, Kano H and Sakaki H 1991 *Appl. Phys. Lett.* **59** 700
- [15] Colas E, Shahar A and Tomlinson W J 1990 *Appl. Phys. Lett.* **56** 955
- [16] Dilger M, Hohenstein M, Phillipp F, Eberl K, Kurtenbach A, Grambow P, Lehmann A, Heitmann D and von Klitzing K 1994 *Semicond. Sci. Technol.* **9** 2258
- [17] Weis J, Haug R J, von Klitzing K and Ploog K 1992 *Phys. Rev. B* **46** 12 837

Non-invasive detection of single-electron processes

M Field, C G Smith, M Pepper, K M Brown, E H Linfield,
M P Grimshaw, D A Ritchie and G A C Jones

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Abstract. Electrical transport measurement can be made 'non-invasively' by capacitively coupling a detector circuit with known behaviour to a quantum device. This allows much more sensitive measurements to be performed; since no electrons are exchanged with the device, the amount of disturbance can be reduced. The sensing circuit is still operational in regimes of extreme low current in the quantum device. This article is a review of work with two non-invasive probes: a voltage probe capable of sensing single-electron effects at currents of order 10^{-17} A, and a capacitance detector which is able to detect the localization of electrons where no electrons move on a timescale of seconds.

1. Introduction

Single-electron effects in both metal in semiconductor systems have been seen and understood for several years now [1]. They were first identified in granular metal samples [2] and have also been observed in semiconductor structures employing 'electrostatic squeezing' [3]. Meirav *et al* [4], and later Kouwenhoven *et al* [5], demonstrated that the conductance of a lateral quantum dot [6] weakly coupled to connecting leads via tunnel barriers oscillates when the potential is changed on a nearby gate. These oscillations, caused by single-electron charging of the quantum dot, have become known as Coulomb blockade (CB) oscillations. These all make use of the large energy associated with charging a small capacitor ($< 10^{-16}$ F) where the energy required to add a single extra electron is larger than the thermal energy at cryogenic temperatures. The conductance of a small island (or 'quantum dot' in semiconductor systems) is governed by these charging effects; current can only flow when extra electrons are allowed to tunnel on and off the island. Since the island becomes charged, the potential of the dot changes as electrons tunnel on and off,

2. The non-invasive voltage probe

We have constructed a device which can detect this change in potential of the dot and is described in detail in [7, 8]. The circuit containing the island is coupled closely to a separate sensing circuit with a single quantum point contact. The point contact is biased so that it becomes a tunnel barrier. The height of the tunnel barrier is dependent on the local electrostatic conditions, and hence the resistance of the detector circuit follows the charging of the dot. The whole system can be calibrated to give a quantitative measurement of the electrostatic potential. This detector

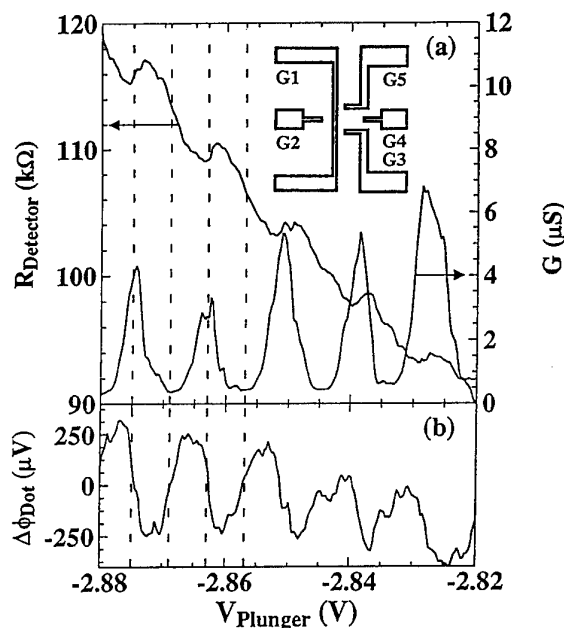


Figure 1. (a) CB oscillations of conductance versus gate voltage through the dot, together with the resistance of the detector circuit. (b) The change in dot potential calculated from the detector resistance. The overall negative slope is an artefact of the calibration procedure. Inset: a schematic diagram of the gate structure.

is a truly non-invasive voltage probe since it does not need to exchange electrons with its surroundings, unlike a conventional laboratory voltmeter. This detector continues to work when the current through the quantum dot is made too small to measure, enabling observation of single-electron charging at currents of order 10^{-17} A.

A diagram of the gate geometry used is shown in the inset to figure 1. The gates are fabricated on top of a GaAs/AlGaAs heterostructure with a two-dimensional electron gas (2DEG) 700 Å below the surface of the semiconductor. The bar down the middle (G1) separates two electrical circuits which interact via the electric fields present between the regions in close proximity. To the right of the bar is a quantum dot, formed by gates G3, G4 and G5, with lithographically defined dimensions of $0.75 \times 0.34 \mu\text{m}$. When the voltage on G4 is swept, CB oscillations are seen in the conductance of the quantum dot with a period of $12.28 \pm 0.46 \text{ mV}$. To the left of the bar is a detector circuit with a channel formed by gates G1 and G2 that is sensitive to the electrostatic potential of the dot. The resistance of the channel is set at $150 \text{ k}\Omega$, where it is most sensitive to the surrounding potential.

Figure 1(a) shows the conductance through the dot and the resistance of the detector as the voltage on G4 is swept. The detector resistance has small dips on a rising background which directly correlate with the CB oscillations in the quantum dot. The detector can be calibrated by applying a voltage to the quantum dot region, after which the detector resistance can be transformed into the dot potential as shown in figure 1(b). The electrostatic potential of the dot $\Delta\phi_{\text{Dot}}$, is oscillatory with an average amplitude of $500 \pm 100 \mu\text{V}$. The charging energy of the dot can also be obtained by measuring the total capacitance of the dot to all other conducting regions around it. The capacitances of each conducting region are added up to give the total capacitance $C_{\Sigma} = (2.92 \pm 0.2) \times 10^{-16} \text{ F}$, which implies a potential on the dot due to charging by a single electron $\Delta\phi_{\text{Dot}} = e/C_{\Sigma} = 550 \pm 30 \mu\text{V}$. This is in good agreement with the value obtained from the detector, especially since the calibration is known to slightly underestimate $\Delta\phi_{\text{Dot}}$.

The peaks and troughs of the oscillations in conductance, indicated by the broken lines, correspond to the mean value of the electrostatic potential oscillations. In a conductance minimum, the Fermi energy in the connecting leads lies approximately midway between the coupled N th electron state and the inaccessible $(N+1)$ th state. As the voltage on the external gate is made more positive, the energy of the dot is reduced with respect to the Fermi energy in the connecting leads. When the $(N+1)$ th electron level comes within the energy range of the electrons in the leads, electrons can tunnel in. The electron energy levels in the dot move up and down by the charging energy. The dot is occupied 50% of the time by an extra electron and therefore the potential takes the mean value. In between these positions the percentage occupation by an extra electron either lags or leads the charge required, causing the oscillation in potential.

The tunnel barrier heights were increased until the conductance oscillations were only just measurable. As the voltage on gate G4 was swept negatively, conductance oscillations were observed, which died away as the barrier heights to the leads increased and the current through the dot became too small to measure; however, the signal on the detector remains. The oscillations of the detector can be followed for a further 110 periods until the gate

G4 reaches a voltage of -4.05 V , after which the signal vanishes (see figure 2). As the oscillations disappear, the background slope of the detector resistance curve becomes much steeper, suggesting that the screening of gate G4 by electrons passing through the dot has changed. The oscillations continue until the current through the dot decreases to less than one electron through the dot in the time for half a period of the applied a.c. bias. Extrapolation of the observed CB peak heights suggests the current at the point where the detector signal fails is of order 10^{-17} A , which correlates well with the a.c. signal of order 100 Hz used in this experiment.

The dot potential, shown in the insets to figure 2, has become a sawtooth waveform. The effective width of the conductance peaks gets smaller and this changes the shape of the potential waveform by reducing the relative width of the peaks with respect to the troughs.

3. Non-invasive capacitance measurement

It is also possible to use the CB in a quantum dot as a very sensitive capacitance meter; this work is described in detail in [9, 10]. Using a double 2DEG sample, the dot can be placed in the upper 2DEG on top of a second 2DEG. The capacitance between a larger gate underneath the sample and the quantum dot can be measured by observing the CB period and changes in capacitance as small as 1 aF can be detected. The 2DEG in between tries to screen the electric field, and the amount of screening is due to the ability of the electrons to rearrange which is dependent on the density of states and the diffusivity of the electrons, allowing the full thermodynamic density of states of the screening 2DEG to be measured. This technique is also used to probe the motion of electrons when the conductance of the sample is many orders of magnitude lower than that which can be measured directly. When the electrons become localized on the timescale of the measurement, the screening fails which allows the diffusivity of electrons to be studied directly.

A double 2DEG sample was grown with two 200 Å quantum wells a distance of 200 Å apart, and a backgate 3300 Å beneath the lower 2DEG. The mobility of the upper layer was $7 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ at a carrier concentration $N_t = 4.5 \times 10^{11} \text{ cm}^{-2}$, while the lower layer had a low mobility of $3 \times 10^4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ at the same carrier concentration, $N_b = N_t$. The carrier concentration in the lower layer was controllable via the backgate, complete depletion occurring at approximately -1.4 V .

The sample was patterned into a thin Hall bar shape ($0.6 \mu\text{m}$ wide by $2 \mu\text{m}$ long) by low-energy gallium ion beam damage [11]. Two independent surface Schottky gates of width approximately 500 Å cross the Hall bar a distance of $0.4 \mu\text{m}$ apart. Using these surface gates a quantum box can be induced in the upper 2DEG. The completed device is shown schematically in figure 3.

We measure the screening non-invasively over a small area using the quantum box in the upper 2DEG layer. At zero magnetic field the period of CB oscillations is seen to change dramatically at a backgate voltage of -1.42 V (figure 4) from 7 mV to 2 mV as the screening layer is depleted. The conductance of the back 2DEG dropped

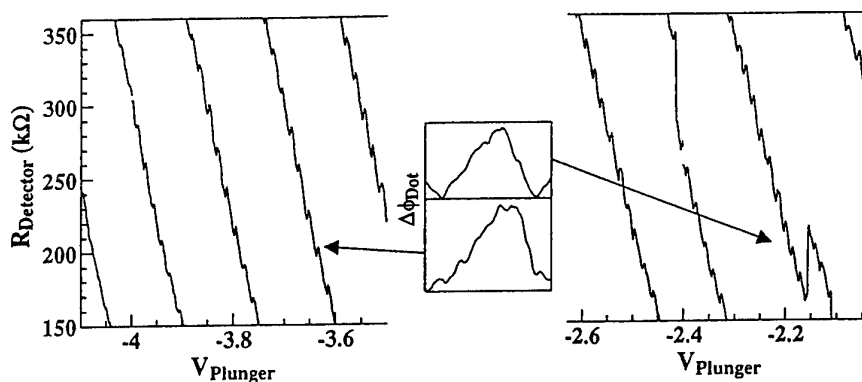


Figure 2. The detector resistance versus V_{G4} after the current through the dot has become too small to be measured. The detector resistance is reset to 150 k Ω whenever it reaches 360 k Ω to maintain sensitivity. The resistance curves follow on from each other. The insets show the inferred dot potential at two points; this has changed from the oscillatory shape seen in figure 1 to a sawtooth shape.

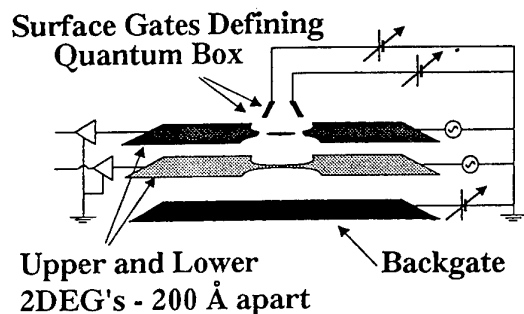


Figure 3. A schematic diagram of the double 2DEG device.

below the sensitivity of the measurement at a more positive gate voltage (-1.3 V). The point of complete depletion can only be observed in the screening; the localization of carriers observed here is the point where the electrons cannot move to screen during the time it takes to sweep the backgate voltage by one CB period, approximately 1 s. As the back 2DEG depletes out it may undergo an Anderson transition and screening then occurs by thermal activation to the mobility edge or by variable range hopping. The actual pinch-off underneath the quantum dot may well occur at different backgate voltages to the observed turn-on of the back 2DEG conductivity. The screening will depend on whether the states underneath the dot are localized, and if not whether they can reach the ohmic contacts which define the ground potential.

4. Summary

In conclusion we have reviewed work on two non-invasive measurement techniques. A non-invasive voltage probe can measure electrostatic potential changes of the order of 50 μ V without exchanging electrons with the device it is measuring. The non-invasive capacitance probe uses a Coulomb blockade electrometer to measure changes in

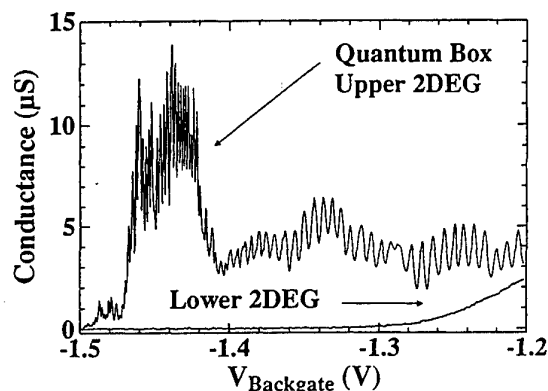


Figure 4. Coulomb blockade oscillations which change period suddenly when the lower 2DEG is fully depleted and can no longer screen. The conductance of the lower 2DEG falls below the minimum sensitivity of the equipment before the 2DEG is completely depleted.

capacitance as low as 1 aF, allowing changes in the screening ability of a 2DEG to be observed.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council and ESPRIT project no. BRA6536.

References

- [1] For a review of Coulomb blockade effects see Grabert H and Devoret M H (ed) 1992 *Single Charge Tunnelling (NATO ASI Series B, vol B294)* (New York: Plenum)
- [2] Giaever I and Zeller H R 1968 *Phys. Rev. Lett.* **20** 1504
Lambe J and Jaklevic R C 1969 *Phys. Rev. Lett.* **22** 1371
- [3] Thornton T J *et al* 1986 *Phys. Rev. Lett.* **56** 1198
- [4] Meirav U *et al* 1990 *Phys. Rev. Lett.* **65** 771
- [5] Kouwenhoven L P *et al* 1991 *Z. Phys. B* **85** 367
- [6] Smith C G *et al* 1988 *J. Phys. C: Solid State Phys.* **21** L893

- [7] Field M, Smith C G, Pepper M, Ritchie D A, Frost J E F, Jones G A C and Hasko D G 1993 *Phys. Rev. Lett.* **70** 1311
- [8] Smith C G, Field M, Pepper M, Ritchie D A, Frost J E F, Jones G A C and Hasko D G 1994 *Surf. Sci.* **305** 553
- [9] Field M, Smith C G, Pepper M, Brown K M, Linfield E H and Jones G A C *Phys. Rev. Lett.* submitted
- [10] Field M, Smith C G, Pepper M, Brown K M, Linfield E H and Jones G A C *Surf. Sci.* at press
- [11] Brown K M *et al* 1994 *J. Vac. Sci. Technol. B* **12** 1293

A possible road to 77 K single-electron devices

S Altmeyer†, F Kühnel, B Spangenberg and H Kurz

Institute of Semiconductor Electronics II, Sommerfeldstraße 24, D-52074 Aachen, Germany

Abstract. The application of single-electron devices is restricted to temperatures at which Coulomb blockade occurs. The most critical parameter of the devices is the tunnelling capacitance. Operation at liquid helium temperature, for example, requires tunnelling capacitances of the order of 20 aF. To realize such low capacitances, advanced electron-beam lithography below 20 nm is mandatory in fabrication schemes proposed up to now. Here we will present a new fabrication process that is capable of tunnelling capacitances as small as 1.5 aF with relaxed technological requirements at the same time. The detailed studies of Coulomb blockades at 20 K currently made may define the route to 77 K operation.

1. Introduction

In recent years the limits of CMOS technology have become evident. A reduction in size to less than 60 nm will not be possible because of the intrinsic size of the p–n junctions. The heat generation in CMOS devices due to the large current densities required for opening and closing down the conducting channel poses an inherent limitation on further reduction. This feature also limits CMOS clock speed. These limitations have initiated an intensive search for alternative low-power concepts. Among them single-electron devices (SEDs), originally proposed by [1], appear very promising. They allow an ultimate reduction in device size down to the atomic scale and the speed of operation is expected to be in the upper GHz range.

2. Principles

The crucial point of SEDs is the restriction in operating temperature. It can only be raised if the device capacitances are lowered appropriately. For the industrial success of SEDs, however, room-temperature operation is inevitable. Therefore most of the experimental effort concerns the fabrication of ultra-small tunnelling capacitances.

Two tunnelling capacitances C_{T1} and C_{T2} in series with a bias voltage source U_B form the backbone of a single-electron transistor (SET). As shown in figure 1, an additional gate capacitance C_G and a gate voltage source U_G are coupled to the island between the tunnel junctions. The main features can be explained by simple considerations of electrostatics.

The electrostatic energy of the charged system is composed of two parts:

$$E = E(n) + E_0.$$

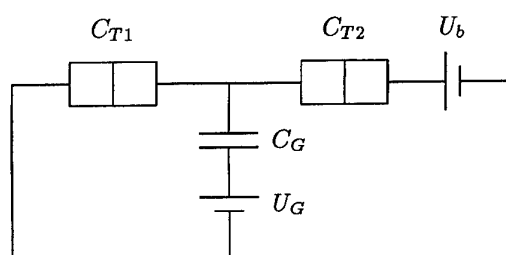


Figure 1. Block diagram of a SET.

$E(n)$ is the energy of the n electrons charging the island and E_0 sums up the outer charges of the capacitances.

$$E(n) = \frac{1}{2} \frac{(n(-e))^2}{C_{T1} + C_{T2} + C_G} + n(-e) \frac{U_B C_{T2} + U_G C_G}{C_{T1} + C_{T2} + C_G}. \quad (1)$$

The charging energy $E(n)$ consists of two terms. The first term describes the n island electrons charging all capacitances connected to the island in parallel, whereas the second term takes into account the potential difference the electrons pass when they enter the island.

From this equation the regime of Coulomb blockade can easily be derived. In the case of

$$E(n-1) \geq E(n) \leq E(n+1) \quad (2)$$

$$E(n-1) - eU_B \geq E(n) \leq E(n+1) + eU_B$$

no electron may enter or leave the island, i.e. no charge transport is allowed. Application of equation (1) leads to

$$\begin{aligned} (n - \frac{1}{2}) \frac{e}{C_G} - U_B \frac{C_{T2}}{C_G} &\leq U_G \leq (n + \frac{1}{2}) \frac{e}{C_G} - U_B \frac{C_{T2}}{C_G} \\ (n - \frac{1}{2}) \frac{e}{C_G} + U_B \frac{C_{T1} + C_G}{C_G} &\leq U_G \\ &\leq (n + \frac{1}{2}) \frac{e}{C_G} + U_B \frac{C_{T1} + C_G}{C_G}. \end{aligned} \quad (3)$$

† Email: altmeyer@basl.rwth-aachen.de.

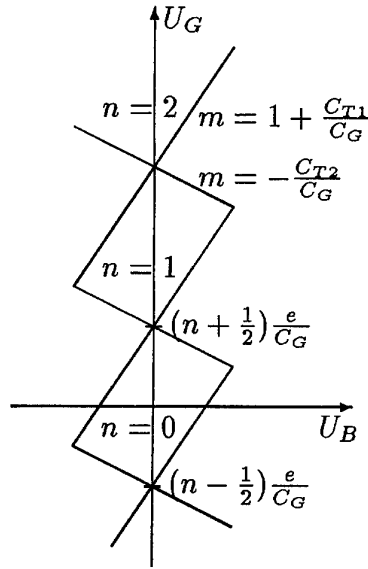


Figure 2. Stability diagram of a SET.

Usually equations (3) are plotted on a U_B - U_G diagram, the stability diagram of a SET, as shown in figure 2.

Above a certain bias voltage U_B , the Coulomb blockade is surpassed and a current is allowed to pass through the device. At a suitable, fixed bias voltage, a gate voltage sweep will cause the system to alternate between the blocking and the conducting states; this is the transistor operation of the device.

The operating temperature of a SED is restricted to temperatures at which the charging energy of a single electron exceeds its thermal energy. Otherwise the Coulomb blockade effects would smear out or would not be observable at all. The maximum allowed operating temperature of a SET can be derived from equation (1):

$$T \leq \frac{e^2}{8k} \frac{1}{C_{T1} + C_{T2} + C_G}. \quad (4)$$

The key point for operation at 77 K is to keep the device capacitances below 1 aF.

3. Fabrication: technology

Many different fabrication schemes for low-capacitance tunnel junctions have been tried.

The oldest fabrication scheme is the two-angle shadow evaporation method [2], which is still frequently used. Typical capacitances achieved with this method are in the range of 50 aF. The self-aligned in-line (SAIL) [3] process is capable of capacitances that are at least half as small. The idea of utilizing thin granular metal films is old, but still a point of intensive research [4–6]. The huge device resistances, however, are still a problem. Recently the local oxidation of thin titanium films with a scanning tunnelling microscope (STM) [7] has been proposed as a possible fabrication scheme for ultra-small tunnelling capacitances. Alternatively the oxidation of prepatterned Si wires for the

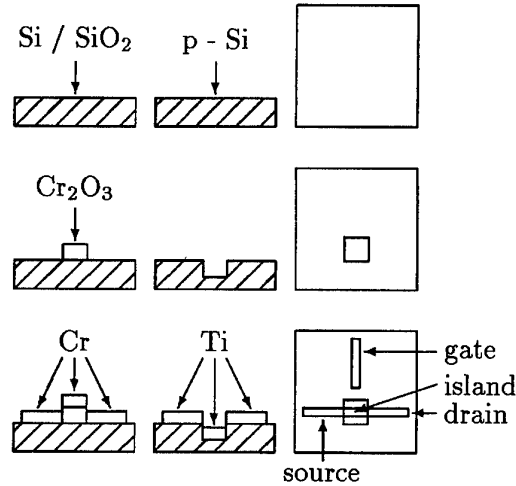


Figure 3. Fabrication scheme of a SECO SET.

formation of nano wires and tunnelling contacts has been reported [8, 9].

An alternative way to meet these challenges with standard electron-beam lithography (EBL) is the step edge cut-off (SECO) [10–13] fabrication scheme, which is also used by other groups [14]. The two possible variants of this process are sketched in figure 3. The additive process is shown on the left of figure 3 and the subtractive fabrication scheme is sketched in the middle. On the right side of figure 3 a top view of both is given.

Using conventional EBL, lift-off and reactive ion etching techniques, a metal line is evaporated either across a recess or a step of a dielectric material on top of the substrate. If the thickness of the metal line is chosen properly with respect to the step height, two interruptions of the metal line in the nanometre range will appear, forming two tunnelling contacts in series, with a metallic island in between. The electric potential of the island can be influenced by an additional gate electrode in close vicinity.

The advantage of this approach is given by the fact that there is no overlap of the metal strips at all, especially not at the edges of the step, which form the tunnelling contacts. This is the reason for the extremely low capacitances of the SECO junctions.

4. Experimental results

Several devices were fabricated using the SECO scheme. Both the additive Cr/Cr₂O₃ and the subtractive p-Si/Ti technologies are suitable for SET production. All samples were electrically characterized at different temperatures, ranging from 30 mK† up to 77 K. Devices with linewidths of 120 nm operate at liquid helium temperatures. The characteristic Coulomb blockade of such a p-Si device is shown in figure 4.

As explained in section 2, the Coulomb blockade shifts under the influence of an applied gate voltage. This shift

† All measurements below $T = 4.2$ K were performed by the PTB Braunschweig.

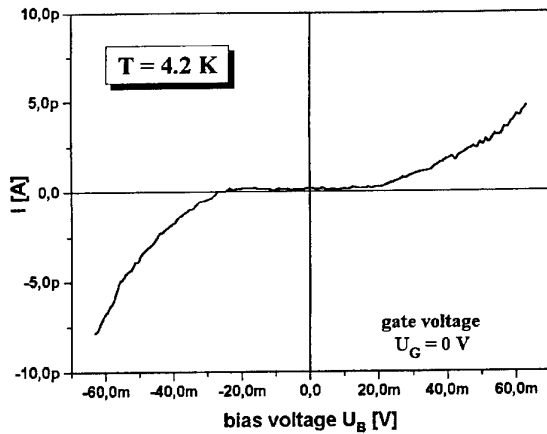


Figure 4. A Coulomb blockade at 4.2 K. The width of the blockade corresponds to a tunnelling capacitance of $C_T = 4$ aF.

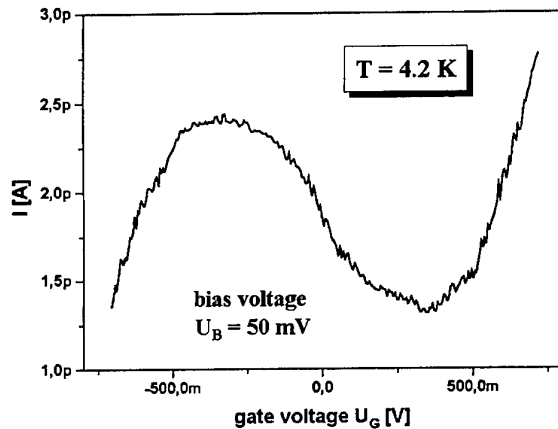


Figure 5. Gate modulation of the output current.

is the basis for a transistor operation and proves that the observed nonlinearities in the I - U characteristics are due to Coulomb blockade effects. Figure 5 shows the experimental observation of this effect. The output current of the SET is plotted as a function of the gate voltage U_G . A current modulation of 1 pA is observed.

If the linewidth is reduced to the range of 75 nm, the tunnelling capacitances are lowered further. As a consequence, a significantly wider Coulomb blockade is observed. As shown in figure 6, the blockade is operative between $U_B = \pm 50$ mV, corresponding to a tunnelling capacitance of $C_T = 1.5$ aF.

These extraordinarily small values for the tunnelling capacitances raise the temperature of operation up to 20 K. Even at liquid nitrogen temperatures, nonlinearities in the I - U characteristics are observable.

The investigated samples showed long-term stability and were robust with respect to repeated thermal cycling. Degradation effects, especially electromigration effects, were not observed. From the geometry of the devices and their electrical characterizations, the current densities in the

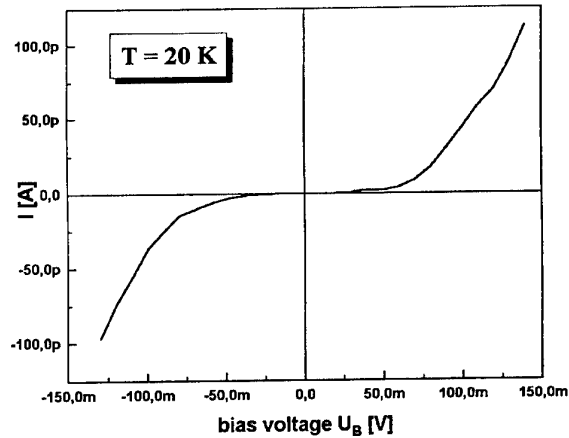


Figure 6. A Coulomb blockade at 20 K.

tunnelling contacts can be calculated to be $j \leq 200$ A cm $^{-2}$. Electromigration effects, however, are known to become observable at typical current densities in the region $j \approx 10^5$ A cm $^{-2}$. Therefore electromigration effects should be no problem for SECO devices, even if the size of the devices was reduced by two orders of magnitude.

5. Conclusion and outlook

The SECO fabrication scheme seems to open the way to 77 K SET operation. It is a simple and robust fabrication scheme for metal SETs with reduced technological requirements, fully compatible with the established silicon technology. Stable 20 K operation was demonstrated and tunnelling capacitances as small as 1.5 aF were realized, using standard EBL. Conventional fabrication processes would require sub 5 nm EBL to achieve this level. A further reduction of the capacitances is possible by using materials with lower dielectric constants and fabricating structures with smaller linewidths. It is safe to assume that a reduction of the existing p-silicon/titanium system to 20 nm linewidth would reduce the capacitances appropriately for 77 K operation. The final goal of stabilized 77 K operation will require further investigations on the reduction of co-tunnelling effects and background charge fluctuations.

Acknowledgments

The authors would like to thank the BMBF for financial support under contract number 13N6256 and the PTB Braunschweig for the mK characterization of our samples.

References

- [1] Averin D V and Likharev K K 1986 *J. Low Temp. Phys.* **62** 345
- [2] Niemeyer J 1974 *PTB-Mitteilungen* **84** 251
- [3] Götz M, Blüthner K, Krech W, Nowack A, Fuchs H-J, Kley E-B, Thieme P, Wagner Th, Eska G, Hecker K and Hegger H 1995 *J. Appl. Phys.* **78** 5499

- [4] Chen W, Ahmed H and Nakazoto 1995 *Appl. Phys. Lett.* **66** 3383
- [5] Chen W and Ahmed H 1995 *J. Vac. Sci. Technol. B* **13** 2883
- [6] Miura N, Yoshikawa N and Sugahara M 1995 *Appl. Phys. Lett.* **67** 3969
- [7] Matsumoto K, Segawa M, Oka Y, Vartanian B J and Harris J S 1996 *Appl. Phys. Lett.* **68** 34
- [8] Leobandung E, Guo L, Wang Y and Chou S Y 1995 *Appl. Phys. Lett.* **67** 938
- [9] Leobandung E, Guo L and Chou Y 1995 *Appl. Phys. Lett.* **67** 2338
- [10] Altmeyer S, Spangenberg B and Kurz H 1995 *Appl. Phys. Lett.* **67** 569
- [11] *German Patent* P 44 41 502.8
- [12] *International Patent* PCT/EP95/04565
- [13] Altmeyer S, Kühnel F, Spangenberg B and Kurz H 1996 *Microelectron. Eng.* **30** 399
- [14] Langheinrich W and Ahmed H 1995 *Japan. J. Appl. Phys.* **34** 6956

Tunnelling through quantum dots

R H Blick, T Schmidt, R Haug and K von Klitzing

Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1,
D-70569 Stuttgart, Germany

Abstract. Tunnelling through single and coupled GaAs quantum dots is used to analyse the electronic properties of artificial atoms and molecules. In addition we demonstrate that a strongly localized dot can be used as a spectrometer with high spatial and energy resolution for the determination of the local density of states in a disordered semiconductor.

1. Introduction

Quantum dots are characterized by a relatively small number of ‘free’ electrons (normally electrons in the conduction band of a semiconductor with an effective mass m^*) confined within an island with submicrometre dimension. For optical experiments small clusters in glass or directly grown dots of InGaAs in GaAs or InP in GaInP (Stranski–Krastanov growth mode in epitaxy) can be used, but for transport measurements a well defined position of the dot relative to the tunnelling barriers is necessary. If the tunnelling resistance is much larger than the quantum resistor $R_K = h/e^2 \approx 26 \text{ k}\Omega$ a well defined integral number of electrons within the dot can be realized. However, the coupling of the dot to the electrical contacts has still to be large enough to allow transport measurements. Recently it has been demonstrated that both the quantum dot and the tunnel barriers can be simultaneously grown by MBE on patterned substrates (see the paper by M Dilger *et al* in this issue) but most of the published data are based on transport measurements on two-dimensional electron systems laterally confined by nanolithography. The quantum dots in all these experiments have the shape of a disc with the thickness of a two-dimensional electron gas (about 10 nm) and a diameter of typically 300 nm. Two types of geometries are normally used for transport measurements which differ mainly by the connection of the ‘quantum disc’ to external reservoirs:

(i) Vertical dot systems, where the tunnel current flows perpendicular to the plane of the disc. Such a device is basically a double-barrier heterostructure similar to a resonant tunnelling diode. The lateral confinement is realized by a mesa etching process. The analysis of the current–voltage characteristic gives information about the charging of the dot (starting with *one* electron), the variation of the ground state energy of the few electron system in a magnetic field, the excited states of the dot system and the local density of states in the emitter and collector contacts. Recent experimental results will be presented in section 2.

(ii) Planar dot system where the ‘quantum disc’ is formed by a negative gate voltage applied to

lithographically patterned split gates on top of a heterostructure with a two-dimensional electron gas (2DEG). The depletion of the 2DEG beneath the gate structures is used for both the lateral confinement of the ‘quantum disc’ and the formation of tunnel junctions connecting the dot to the undisturbed two-dimensional electron gas. The tunnelling barriers can be continuously changed by varying the gate voltage so that for example the coupling between adjacent dots can be adjusted in such a way that the transition from two quantum dots to one quantum molecule can be studied. In addition, capacitively coupled electrodes allow an adjustment of the electrostatic energy of the dots so that single-electron transport is possible even in the limit of small bias voltages across the structure. The planar dot system is therefore much more flexible than the vertical structure. However, contrary to the vertical system where only one electron in a potential minimum within the disc area leads to a tunnel current, the planar structure is not useful for the study of few-electron systems. The lateral extension of the few-electron wavefunction is normally much smaller than the disc area formed by the gates, so that the transmission through the relatively thick tunnelling barriers becomes unmeasurably small. Therefore an accurate determination of the *number* of electrons in a lateral quantum dot is difficult. In section 3 recent experiments on a planar double-dot system with a relatively large number of electrons ($N > 30$) are presented.

Transport measurements on quantum dots at low temperatures are dominated by the Coulomb charging energy. The energy necessary to add one electron onto an island with a capacitance C is $\Delta E = e^2/C$. The charging energy for a GaAs island with a diameter of $0.4 \mu\text{m}$ is of the order of 1 meV, whereas the level separation due to size quantization is one order of magnitude smaller. A detailed introduction to single-charge tunnelling and an overview about theoretical and experimental work in this field can be found in the proceedings of different conferences [1–3]. In the following sections we present some special aspects of experiments on vertical and lateral GaAs dots.

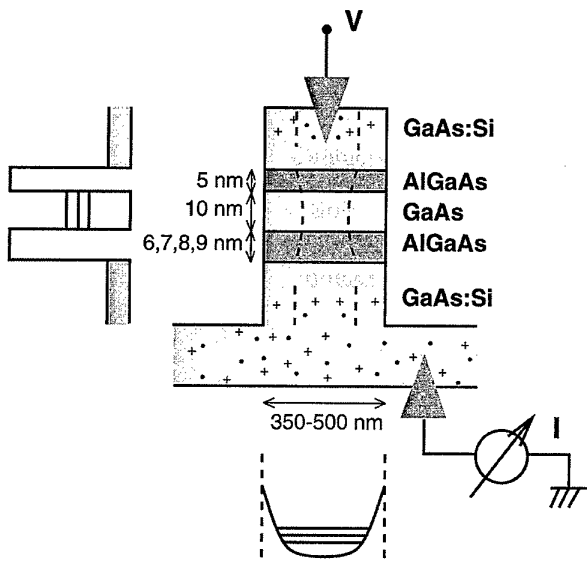


Figure 1. Cross section of a double-barrier tunnelling device used for single-electron spectroscopy.

2. Single-electron tunnelling in vertical dot systems

Ashoori *et al* have shown [4] that a laterally confined 2DEG can be charged with a discrete number N of electrons ($0 < N < 35$) by tunnelling from the substrate through a depletion layer into the dot. The change of the dot charge leads to a charging current which can be measured directly. Such an experiment gives information about the energies of the N -electron ground state. Similar results are obtained if the 'leakage' current of a charged dot is measured. Such a configuration is shown in figure 1. The dot is charged through a thin emitter-dot tunnel barrier of 5 nm thickness (negative potential at the top electrode). The 'leakage' current I through the relatively thick dot-collector barrier (for example, 9 nm width) is directly proportional to the number of electrons in the dot (if the tunnel probability through the thick barrier remains constant in the range of applied emitter-collector voltages).

The device structure shown in figure 1 is used for all experiments presented in this section. Details about the growth and fabrication of the devices have been published in a previous paper where the same structures were used to analyse the peak-to-valley ratio of small resonant tunnelling devices [5].

In the following, transport measurements on small double-barrier tunnel devices are presented where two parameters are varied, the diameter of the mesa structure ('small' or 'large' devices) and the direction of the current through the asymmetric structure. The notation 'small' device will be used if the geometrical diameter of the device is smaller than 350 nm and the lateral confinement is dominated by the nearly parabolic depletion potential as sketched in figure 1. (It should be noted that the electrical diameter of the dot, especially for the few-electron system, is much smaller than the geometrical diameter.) For 'large' devices not the geometrical dimension but the

potential fluctuations due to impurities dominate the lateral confinement for the first electrons in the dot. In order to avoid too many potential minima with nearly the same energy within the area of the mesa structure (which leads to a large number of impurity-dots in parallel) a finite (but not too small) diameter of typically $0.5\text{--}2\text{ }\mu\text{m}$ is required for the mesa structure.

The simplest current-voltage characteristic, which corresponds to the capacitance measurements in [4] is observed for 'small' devices under charging conditions where the electrons are injected through the thin emitter-dot barrier and are accumulated in the dot due to the small transmission of the thick dot-collector barrier. Figure 2 shows a typical result with the current-voltage characteristic in figure 2(a) and a grey-scale plot of the differential conductance dI/dV as a function of the magnetic field B in figure 2(b). The magnetic field orientation is parallel to the current direction. The number of electrons in the dot increases step-like with increasing bias voltage V . Every time the electrochemical potential of the dot $\mu(N) = E(N) - E(N-1)$ ($E(N)$ is the ground state energy of the N -electron dot) crosses the Fermi energy of the emitter, the dot charge changes by one electron. The step height for the current in figure 2(a) is dominated by the transmission probability of electrons through the thick collector barrier. Systematic studies on devices with different collector barrier thicknesses show that the step height decreases exponentially if the barrier thickness is increased [6]. The analysis of the magnetic-field-dependent charging process (figure 2(b)) gives information about the ground state energy of quantum dot hydrogen, quantum dot helium, etc. The monotonic increase in the energy of the one-electron system with magnetic field (lowest line in figure 2(b)) contains information about the diameter of the artificial hydrogen (diamagnetic shift) whereas the non-monotonic variation of the other lines can be interpreted as transitions in the ground state energy between states of different angular momentum and spin (for example spin-singlet to triplet transitions for quantum dot helium in a magnetic field [7]).

A similar device as used for the measurements in figure 2, but with an inverted polarity for the bias voltage (non-charging condition, injection of electrons into the dot through the thick barrier), shows a quite different behaviour (figure 3). Fine structure is visible in the plateau region due to the injection of electrons into excited dot states $\mu^*(1)$. This process influences the current since the thick emitter-dot barrier dominates the total transmission of electron through the double-barrier structure. The grey-scale plot in figure 3 shows clearly that the magnetic field dependence of the excited states follows the single-particle energies of one electron in a parabolic confining potential. A detailed analysis of excited dot states has been published by Su *et al* [8].

A quite different type of fine structure in the current-voltage characteristic is visible in the non-charging regime of 'large' devices (figure 4). For completeness a reference measurement on the same device as shown in figure 4, but under reversed bias voltage (charging regime where the electrons are injected through the thin tunnel barrier), is

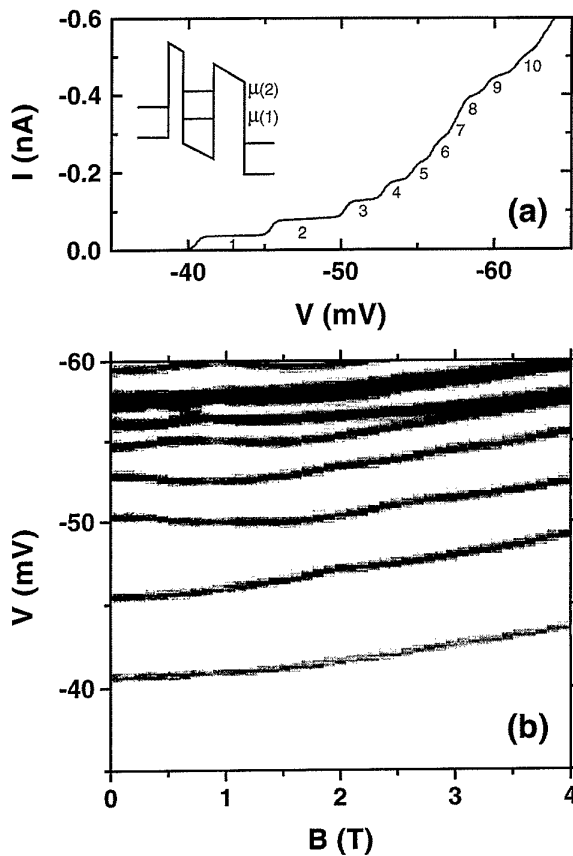


Figure 2. Typical result for the current-voltage characteristics (a) and the magnetic-field-dependent differential conductance dI/dV (in grey scale) (b) of a small dot ($d_{\text{geo}} \approx 350$ nm) in the charging regime.

shown in figure 5. The current-voltage characteristic shows a step-like behaviour similar to the curve in figure 2(a), but with a larger current since the thickness of the collector barrier is 8 nm compared with 9 nm for the small device in figure 2. The weak magnetic field dependence of the one-electron energy indicates that the first electron in the dot is strongly confined (on a state of about 20 nm diameter at a geometrical diameter of $0.5 \mu\text{m}$ for the mesa structure). The parallel shift of the second line in figure 5(b) indicates that the second electron is not used to form quantum dot helium but more probably a second quantum dot hydrogen spatially separated from the first electron. The strong confinement of the first electron in a large quantum dot seems to be necessary for the observation of the new type of fine structure shown in figure 4.

Basically, fine structure in the non-charging regime of the quantum dot may arise either from the opening of *new* transmission channels (if for example excited dot states cross the Fermi energy of the emitter, see figure 3) or from a fluctuation in the local density of states *below* the Fermi energy of the emitter (see inset in figure 4(a)). In this case the quantum dot acts as a spectrometer with high spatial and spectral resolution. The fluctuations in the local density of states of the emitter in strong magnetic fields ($B > 4$ T) seem to be related to the Landau levels in a

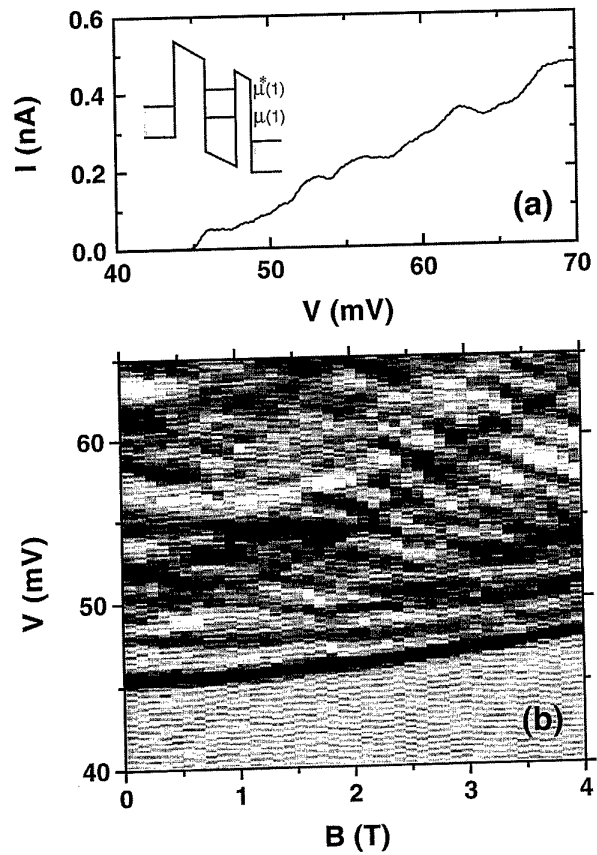


Figure 3. Same experiments as shown in figure 2 but for a bias voltage polarity corresponding to the non-charging regime (see inset of figure 3(a)).

three-dimensional system

$$E = (n + \frac{1}{2})\hbar\omega_c + \frac{\hbar^2 k_z^2}{2m^*} \quad (1)$$

where n is the Landau quantum number, $\hbar\omega_c = \hbar eB/m^*$ is the cyclotron energy and k_z is the wavevector in the direction of the magnetic field.

The line $V_1(B)$ in figure 4(b) which characterizes the resonance of the Fermi energy E_F in the emitter with the single-electron energy $\mu(1)$ of the dot fulfils the equation

$$E_F = \mu(1) - \alpha e V_1(B) \quad (2)$$

where α is the voltage-energy conversion coefficient with $\alpha \approx 0.5$ for the dot energies. In a similar way the fine structure lines $V(B)$ are obtained if the Fermi energy in equation (2) is replaced by the Landau energy E in (1)

$$E = \mu(1) - \alpha e V(B). \quad (3)$$

The combination of (1)–(3) leads to the relation

$$V(B) = V_1(B) + \frac{1}{\alpha e} \left(E_F - \frac{\hbar^2 k_z^2}{2m^*} \right) - \frac{1}{\alpha e} (n + \frac{1}{2}) \frac{\hbar e B}{m^*}. \quad (4)$$

The experimentally observed negative slope dV/dB of the fine structure seems to be related to the term

$$\frac{1}{\alpha e} (n + \frac{1}{2}) \frac{\hbar e}{m^*}$$

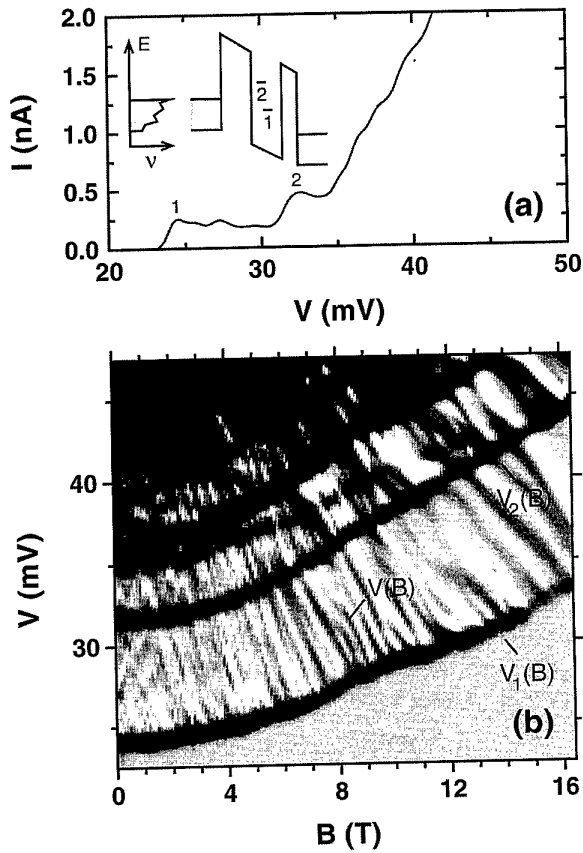


Figure 4. Fine structure in the current-voltage characteristic of a large dot ($d_{\text{geo}} \approx 0.5 \mu\text{m}$) in the non-charging current direction: (a) current-voltage characteristic (the electron flow relative to the thin and thick tunnelling barriers is shown in the inset); (b) grey-scale plot for the differential conductance dI/dV as a function of bias voltage and magnetic field.

(see equation (4)) which contains no adjustable parameter. The agreement of this slope with the experimental data in the magnetic field range around $B = 8 \text{ T}$ where two occupied Landau levels in the emitter are expected ($n = 1$) may be accidental since the slope does not change by a factor of three in higher magnetic fields with $n = 0$. The large number of nearly parallel lines in figure 4(b) can be explained if quantized values for k_z are assumed. We believe that the distribution of impurities in the emitter selects the special k_z values.

In summary, the single-electron tunnelling in double-barrier heterostructures can be used to investigate the N -electron ground state energies of a quantum dot, the excited single-electron state energies and the local density of states in the emitter.

3. Single-electron tunnelling in lateral double quantum dot systems

Most of the published data concerning single-electron transport in semiconductors are based on quantum dots which are formed by negatively biased split gates on top of a GaAs/AlGaAs heterostructure. The main advantage of

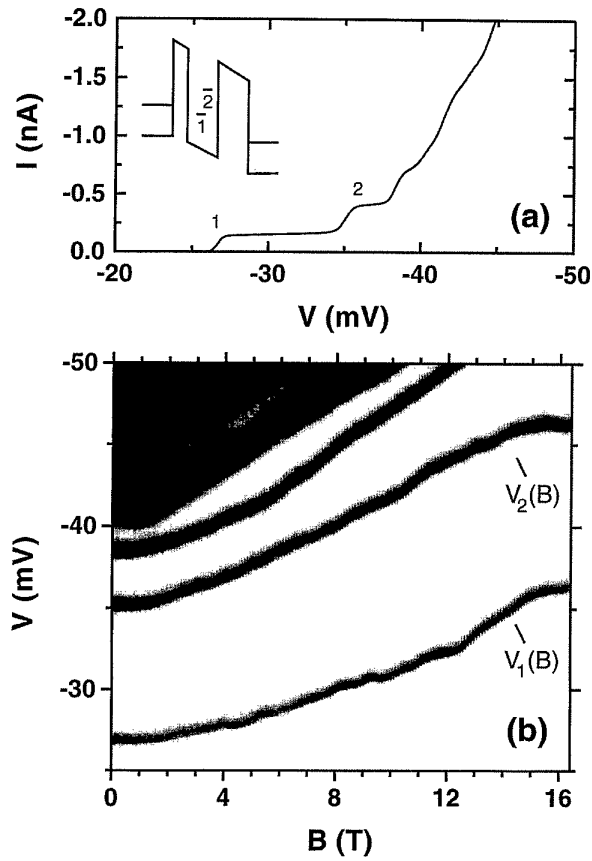


Figure 5. Experimental data for the large dot in the charging regime. (a), (b) as figure 4.

such devices compared with the vertical tunnel structure is that the starting material is a high-quality two-dimensional electron gas and that the lateral confinement does not introduce additional impurities (whereas the high doping of the bottom contact in the vertical structure often leads to a contamination of the dot area with impurities). In addition the electrostatic potential of the lateral dot can be varied with back gate (V_B) and top gate (V_T) voltages. This means that the dot can be charged even without a finite emitter-collector voltage.

At gate voltages V_B, V_T where the dot charge changes by one electron, an emitter-collector current through the dot is possible. In a V_B - V_T diagram, this condition is characterized by parallel lines which separate the Coulomb blockade regimes with a fixed number of electrons in the dot. The slope of the lines is determined by the capacitive coupling of the dot to the gate electrodes. (The lines broaden proportional to the emitter-collector voltage V_{ec} and contain fine structure due to excited states in the quantum dot [9]. This nonlinear behaviour will not be discussed in the following, and only data with excitation voltages V_{ec} less than $5 \mu\text{V}$ will be presented.)

If two independent dots are connected in series, a finite conductivity through the double-dot structure is expected only at gate voltages V_B, V_T where *both* dots have a finite conductivity. This corresponds to the crossing points of two sets of parallel lines in the V_B - V_T charging diagram as shown in figure 6(a). The different slopes of the lines

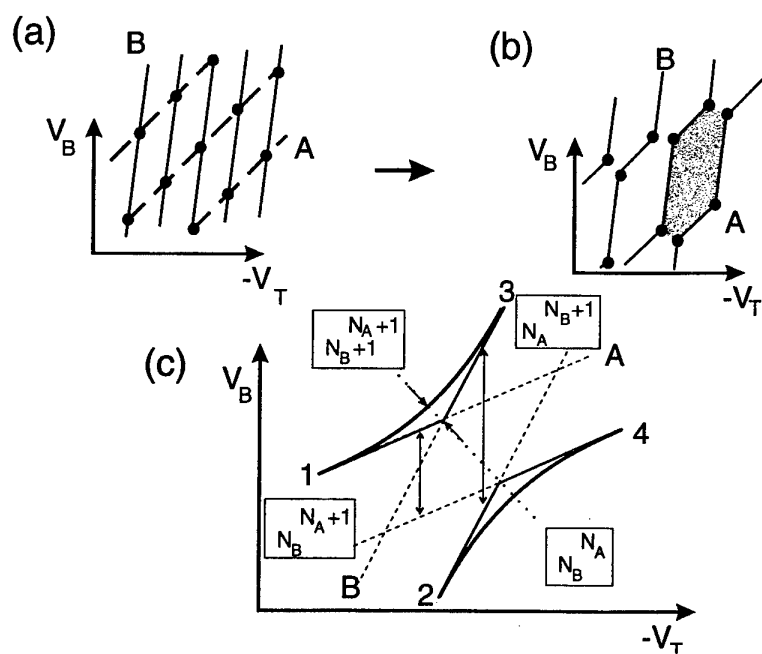


Figure 6. Charging diagram for a double-dot system (DDS): (a) two independent dots; (b) two dots coupled by Coulomb interaction; (c) two dots coupled by both Coulomb interaction and tunnelling processes. The full lines separate the gate voltage regions where the total number of electrons in both dots changes by one elementary charge. The arrows indicate the parallel shift of the lines due to Coulomb interaction and the shift due to tunnelling processes between the two dots.

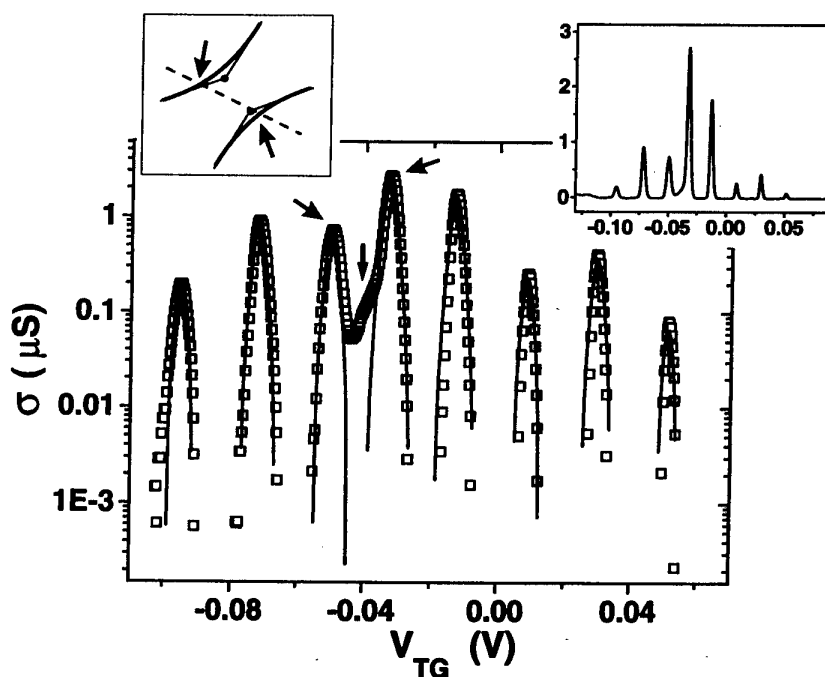


Figure 7. Measured conductivity of a double-dot system as a function of the top gate voltage V_{TG} . The gate voltage variation corresponds to a measurement along the broken line of the charging diagram (see inset). The full lines for the conductivity are calculated including the Coulomb interaction and a finite temperature. The structure at $V_{TG} \approx -0.04$ V originates from the finite tunnelling probability between the two dots.

originate from different capacitive coupling of the dots to the gate electrode, which is the case if the dots A and B have different sizes. The degeneracy at the crossing

points (where the number of electrons N_A and N_B in the two dots can be changed simultaneously) is lifted if the Coulomb interaction between the two dots (characterized

by an interdot capacitance) is included. A variation in the charge of dot A leads to a shift in the electrostatic potential of dot B and vice versa. This results in a parallel shift of the lines in figure 6(a) as shown in figure 6(b). The zig-zag curves separate the areas of the charging diagram which are characterized by a constant *total* charge of *both* dots.

If the separation between the two dots disappears (combination of the two dots to one larger dot), the zig-zags in figure 6(b) become straight lines as expected for a single dot. Up to now the contribution of a finite tunnelling probability between the two dots has not been included and the following discussion concentrates on the influence of tunnelling processes between the dots on the transport characteristic of a double dot system (DDS).

The influence of electron tunnelling between two dots on the charging diagram is shown in figure 6(c). The straight lines correspond to the situation of vanishing overlap of the electron wavefunctions in the two dots (but including Coulomb interaction between the dots) and the curves are expected if the overlap leads to a bonding–antibonding splitting of the ground state of the double-dot system. The existence of a coherent mode in the DDS leads to a finite conductivity through the system even if the gate voltages do not match exactly the positions where sequential single-electron transport through the DDS is allowed (crossing points in the charging diagrams of figures 6(a) and (b)). Basically a finite temperature may lead to a broadening of the singularities for the conductivity in the V_B – V_T diagram but an analysis of the conductivity as a function of both V_B and V_T shows that the so-called ‘stochastic Coulomb blockade’ cannot explain the experimental results. In addition the following results support the idea that a molecular-like state in the DDS can be formed:

(i) Temperature-dependent measurements ($23\text{ mK} < T < 1.6\text{ K}$) through a DDS (for device parameters see [10]) show that the conductivity decreases with increasing temperature in the region where coherent transport is expected. This may be explained by a destruction of the coherence with increasing temperature.

(ii) The coherent transport through a DDS may be destroyed by an emitter–collector voltage larger than the bonding–antibonding splitting (similar to the Wannier–Stark localization in superlattices). Such negative differential conductance has been observed [11].

(iii) As shown in figure 7, additional structures are observed (indicated by a vertical arrow) close to gate

voltages of the main resonance (= largest peak in the logarithmic plot). The corresponding gate voltage sweep within the charging diagram is shown as a broken line in the inset. The variation in the period of the oscillations by about 30% is a result of the dot–dot interaction and the additional structure at $V_{TG} = -0.04\text{ V}$ can be well interpreted as a state resulting from the tunnel splitting. The full lines in figure 7 are the expected lineshapes if a thermal broadening is included.

All these experiments indicate that not only the Coulomb interaction between two dots, which has been discussed by different groups for both metallic [12] and semiconductor systems [13–16], but also the finite tunnelling probability between two dots influences the electronic properties of double-dot systems.

References

- [1] 1991 Special issue on single charge tunneling *Z. Phys. B* **85** 317–467
- [2] 1992 *Single-Electron Tunneling and Mesoscopic Devices* (Springer Series in Electronics and Photonics vol 31) (Berlin: Springer)
- [3] 1993 *Proc. NATO ARW on Physics of Few-Electron Nanostructures* (Physica B **189**)
- [4] Ashoori R C, Stormer H L, Weiner J S, Pfeiffer L N, Baldwin K W and West K W 1994 *Surf. Sci.* **305** 558–65
- [5] Schmidt T, Tewordt M, Haug R J, von Klitzing K, Schönherr B, Grambow P, Förster A and Lüth H 1996 *Appl. Phys. Lett.* **68** 838
- [6] Schmidt T, Haug R J, von Klitzing K, Förster A and Lüth H 1996 *Phys. Rev. B* at press
- [7] Pfannkuche D, Gerhardt R, Maksym P and Gudmundsson V 1993 *Physica B* **189** 6–15
- [8] Su B, Goldman V J and Cunningham J E 1992 *Phys. Rev. B* **46** 7644
- [9] Weis J, Haug R J, von Klitzing K and Ploog K 1994 *Semicond. Sci. Technol.* **9** 1890–6
- [10] Blick R H, Haug R J, Weis J, Pfannkuche D, von Klitzing K and Eberl K 1996 *Phys. Rev. B* **53** 7899
- [11] Blick R H unpublished
- [12] Pothier H, Lafarge P, Urbina C, Esteve D and Devoret M H 1992 *Europhys. Lett.* **17** 249
- [13] Ruzin I M, Chandrasekhar V, Levin E I and Glazman L I 1992 *Phys. Rev. B* **45** 13 469
- [14] Kemerink M and Molenkamp L W 1994 *Appl. Phys. Lett.* **65** 1012
- [15] Waugh F R, Berry M J, Mar D J, Westervelt R M, Campman K C and Gossard A C 1996 *Phys. Rev. B* **53** 1413
- [16] Hofmann F, Heinzel T, Wharam D A, Kotthaus J P, Böhm G, Klein W, Tränkle G and Weimann G 1995 *Phys. Rev. B* **51** 13 872

Photon-assisted tunnelling through a quantum dot

T H Oosterkamp, L P Kouwenhoven, A E A Koolen,
N C van der Vaart and C J P M Harmans

Department of Applied Physics, Delft University of Technology, PO Box 5046,
2600 GA Delft, The Netherlands

Abstract. We report photon-assisted tunnelling (PAT) through a quantum dot with zero-dimensional (0D) states. PAT allows electrons to reach previously inaccessible energy states by absorbing or emitting photons from a microwave signal. We discuss a model based on a master equation for a quantum dot with 0D states and include PAT processes. Simulations are compared with measurements.

1. Introduction

Transport through a quantum dot is regulated by Coulomb blockade effects. The energy to put an extra electron on a quantum dot includes a charging energy E_c and a finite energy difference $\Delta\epsilon$ arising from the confinement. A dot is said to have 0D (zero-dimensional) states if $\Delta\epsilon$ is larger than the thermal energy $k_B T$. Usually, *elastic* tunnel processes through such discrete quantum states are investigated. In the presence of a high-frequency signal *inelastic* transport processes are also possible. Electrons can then reach previously inaccessible energy states by absorbing or emitting photons from the high-frequency signal during tunnelling. This is known as photon-assisted tunnelling (PAT).

The theory of PAT for a single tunnel junction dates back to the work of Tien and Gordon in 1963 [1]. Experiments with high-frequency signals applied to a quantum point contact were reported in [2]. In these experiments, however, PAT could not be demonstrated unambiguously. PAT has been demonstrated in superlattices irradiated by a free electron laser [3]. PAT was also shown in quantum dots without 0D states [4, 5]. Here, we discuss a Tien–Gordon model for quantum dots containing 0D states. We present simulations on the interplay between PAT processes and 0D states. (Such tunnel processes are sketched in the inset of figure 2.) The simulations can be used to study the effects of relaxation processes. Finally, we discuss measurements on PAT through 0D states.

2. PAT through a single junction

Our first step is to look at photon-assisted transport through a single tunnel junction. Suppose that we have an oscillating potential difference across a tunnel junction $\tilde{V} \cos(\omega t)$: in which \tilde{V} is the a.c. amplitude and ω is the angular frequency. This gives a Hamiltonian $H = H_0 + H_{\text{int}} = H_0 + e\tilde{V} \cos(\omega t)$. H_0 is the unperturbed

Hamiltonian describing the two leads on either side of the tunnel junction. The effect of the oscillating potential is that the time-dependent part of the electron wavefunction, when expanded into a power series, contains the energy components E , $E \pm \hbar\omega$, $E \pm 2\hbar\omega$, ..., etc where $\hbar\omega$ is the photon energy. These energy components are called sidebands. The expansion can be done as [1]

$$\begin{aligned}\psi(\mathbf{r}, t) &= \varphi(\mathbf{r}) \exp\left(-i \int dt [E + e\tilde{V} \cos(\omega t)]/\hbar\right) \\ &= \varphi(\mathbf{r}) \exp(-iEt/\hbar) \sum_{n=-\infty}^{\infty} J_n(e\tilde{V}/\hbar\omega) \exp(-in\omega t) \\ &= \varphi\left(\sum_{n=-\infty}^{\infty} J_n(e\tilde{V}/\hbar\omega) \exp[-i(E + n\hbar\omega)t/\hbar]\right) \quad (1)\end{aligned}$$

$\varphi(\mathbf{r})$ is the space-dependent part of the wavefunction $\psi(\mathbf{r}, t)$. $J_n(\alpha)$ is the n th-order Bessel function of the first kind evaluated at $\alpha = e\tilde{V}/\hbar\omega$. The sidebands are only well defined if the number of tunnel events per unit of time (i.e. the tunnel rates) are much smaller than the photon frequency. Because there is no electric field in the scattering-free leads, mixing of electron states in the leads is absent [6]. The probability for tunnelling from an occupied state E to an unoccupied state $E + n\hbar\omega$ is given by $P(E \rightarrow E + n\hbar\omega) = J_n^2(\alpha)$. A positive (negative) n corresponds to the absorption (emission) of n photons during the tunnel process. Elastic tunnelling without photons corresponds to $n = 0$.

A net current flows by introducing an asymmetry, for instance, by applying a d.c. bias voltage V_0 between the two leads. If tunnelling is a weak perturbation, the current in the presence of microwaves \tilde{V} is given by [1]

$$\begin{aligned}\tilde{I}(V_0) &= c \sum_{n=-\infty}^{\infty} J_n^2(e\tilde{V}/\hbar\omega) \\ &\times \int_{-\infty}^{\infty} [f_1(E - eV_0) - f_1(E + n\hbar\omega)] \rho_A(E - eV_0) \\ &\times \rho_B(E + n\hbar\omega) dE\end{aligned}$$

$$= \sum_{n=-\infty}^{\infty} J_n^2(e\tilde{V}/\hbar\omega)I(V_0 + n\hbar\omega/e). \quad (2)$$

$f(E)$ is the Fermi function, ρ_A and ρ_B are the unperturbed densities of states of the two leads, c is a constant, which is proportional to the tunnel conductance, and I is the tunnel current without an oscillating field. For a *single* junction the d.c. current in the presence of microwaves is thus simply described in terms of the d.c. current without microwaves. Note that for a *double* junction equation (2) is not valid, as will be shown in the following.

3. The 0D model based on a master equation

Our model is based on a master equation approach [7] with the inclusion of PAT. We assume $E_c \gg \Delta\epsilon, k_B T, \hbar f = \hbar\omega$ such that we can use a two-state (N and $N+1$) model [8]. We neglect level broadening due to a finite lifetime of the electrons on the dot. A quantum dot state can be described by the total number of electrons on the dot (which is either N or $N+1$) together with χ which describes the particular distribution of electrons over the 0D states. The net current through the dot I follows from the probability $P_{N,\chi}$ that a particular distribution is occupied together with the tunnel rates through one of the barriers:

$$I = \sum_{\chi} \sum_{j=\text{empty}} P_{N,\chi} \Gamma_{l,j}^{\text{in}} - \sum_{\chi'} \sum_{j=\text{full}} P_{N+1,\chi'} \Gamma_{l,j}^{\text{out}} \quad (3)$$

Here $\Gamma_{l,j}^{\text{in(out)}}$ are the rates into or out of 0D state j , through the left barrier:

$$\Gamma_{l,j}^{\text{in}}(\epsilon_j) = \Gamma_1 \sum_n J_n^2(\alpha_l) f(\epsilon_j - n\hbar\omega + eV_0)$$

$$\Gamma_{l,j}^{\text{out}}(\epsilon_j) = \Gamma_1 \sum_n J_n^2(\alpha_l) [1 - f(\epsilon_j - n\hbar\omega + eV_0)] \quad (4)$$

where Γ_1 is the tunnel rate of the left barrier determined by the barrier shape, and ϵ_j is the energy of 0D state j measured relative to the Fermi energy of the right lead. An equivalent set of equations can be given for the right barrier taking $V_0 = 0$. In the following simulations we take equal a.c. amplitudes dropping across the left and right barriers; i.e. $\alpha_l = \alpha_r = \alpha$. The probabilities $P_{N,\chi}$, are calculated from the set of master equations given by

$$\begin{aligned} \dot{P}_{N,\chi} = & \sum_{\chi'} P_{N+1,\chi'} (\Gamma_{l,\chi'}^{\text{out}} + \Gamma_{r,\chi'}^{\text{out}}) \\ & - P_{N,\chi} \sum_{j=\text{empty}} (\Gamma_{l,j}^{\text{in}} + \Gamma_{r,j}^{\text{in}}) + \sum_{\chi'' \neq \chi} P_{N,\chi''} \Gamma_{\chi'' \rightarrow \chi} \\ & - P_{N,\chi} \sum_{\chi''' \neq \chi} \Gamma_{\chi \rightarrow \chi'''} \end{aligned} \quad (5)$$

and its equivalent for $\dot{P}_{N+1,\chi'}$. Note that for $N = 2$ we have ten different χ if we include five different 0D states, so we have ten equations $\dot{P}_{N,\chi}$, and another ten equations $\dot{P}_{N+1,\chi'}$. The first two terms in equation (5) correspond to a change in the occupation probability of a certain distribution due to tunnelling (the number of electrons on the dots changes). The first term takes into account the rates that correspond to an electron tunnelling out of state $j_{\chi'}$ leaving the dot in the distribution (N, χ) . For an electron tunnelling into the dot

one needs to sum over all the states j that are empty when the dot is in state χ (second term). In the last two terms the number of electrons on the dot stays the same, only the distribution over the states changes (i.e. electrons relax to a lower level or they are excited to a higher level). We take excitation rates equal to zero (no mixing of electron states in the dot due to the high frequency) and non-zero relaxation rates.

To find a stationary solution for the occupation probabilities in the dot these equations are set to zero and solved with the boundary condition:

$$\sum_{\chi} P_{N,\chi} + \sum_{\chi'} P_{N+1,\chi'} = 1.$$

The current is calculated for a fixed set of 0D energies $\{\epsilon_j\}$. The effect of a gate voltage is simulated by shifting the 0D energies ϵ_j relative to the Fermi energy of the right reservoir.

Figure 1 shows simulations without relaxation between the states in the dot. We have taken $\Delta\epsilon = 3\hbar\omega$ for the curves in the inset. Next to the main resonance we see that side peaks develop at multiple values of $\hbar\omega/e$ when α is increased. In the main figure we have taken $\Delta\epsilon = 0.75\hbar\omega$. Here not only do side peaks develop but we also see peaks at other gate voltages. These peaks arise due to the interplay between the 0D states and the photon energy. Their locations are described by $(m\Delta\epsilon + n\hbar\omega)/e$ where $m = 0, \pm 1, \pm 2, \dots$ and n is the photon number. Similar simulation results have been reported by Bruder and Schoeller [9].

Figure 2 shows an expansion for the curve $\alpha = 1$. We have assigned the excited states ϵ_j and the particular PAT process. The $N+1$ ground state is denoted by $j = 0$, positive j are excited states above ϵ_0 and negative j are below ϵ_0 . Note that a curve like this resembles an atomic spectrum. The inset shows the effects of relaxation. It is seen that upon increasing the relaxation rate the peaks that correspond to transitions through excited states decrease while the peaks corresponding to transitions through the ground state increase.

4. Experiments

In figure 3 data are shown from a dot with a measured 0D splitting of approximately $140 \mu\text{eV}$. This is four or five times larger than the thermal energy. With a split gate technique a small dot of lithographic size $600 \times 300 \text{ nm}^2$ was formed in the 2DEG of a GaAs/AlGaAs heterostructure with mobility $2.3 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ and electron density $1.9 \times 10^{15} \text{ m}^{-2}$ at 4.2 K . In addition to the d.c. voltages on the gates a microwave signal (0 – 75 GHz) was capacitatively coupled to one of the gates. The effective electron temperature was found to be about 300 mK . The charging energy $E_c = 1.2 \pm 0.1 \text{ meV}$. Using finite bias voltage measurements the energy splitting between the ground state and the first excited state was determined to be approximately $140 \mu\text{eV}$ (four or five times larger than the thermal energy) and the conversion factor from gate voltage to energy could be deduced.

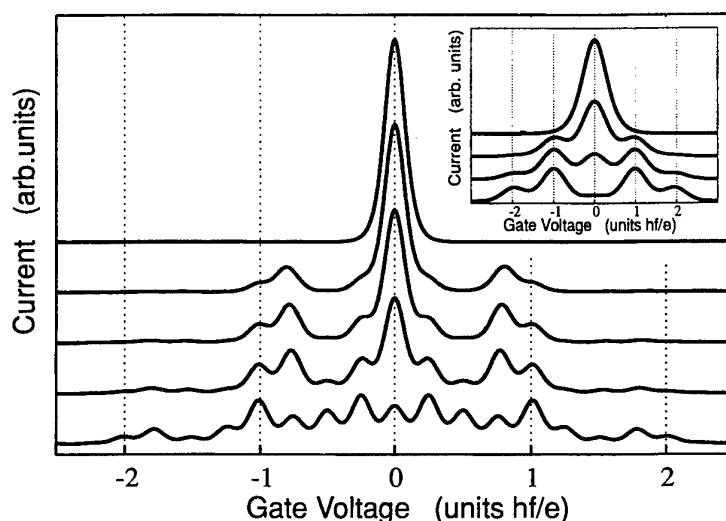


Figure 1. Simulation without relaxation. The parameters for the data in the inset are $\Delta\epsilon = 3\hbar\omega$, $\hbar f = 5k_B T$, and from top to bottom $\alpha = 0, 1, 1.5, 2$. The parameters for the main figure are $\Delta\epsilon = 0.75\hbar\omega$, $\hbar f = 20k_B T$ and from top to bottom $\alpha = 0, 0.5, 0.75, 1, 1.5$.

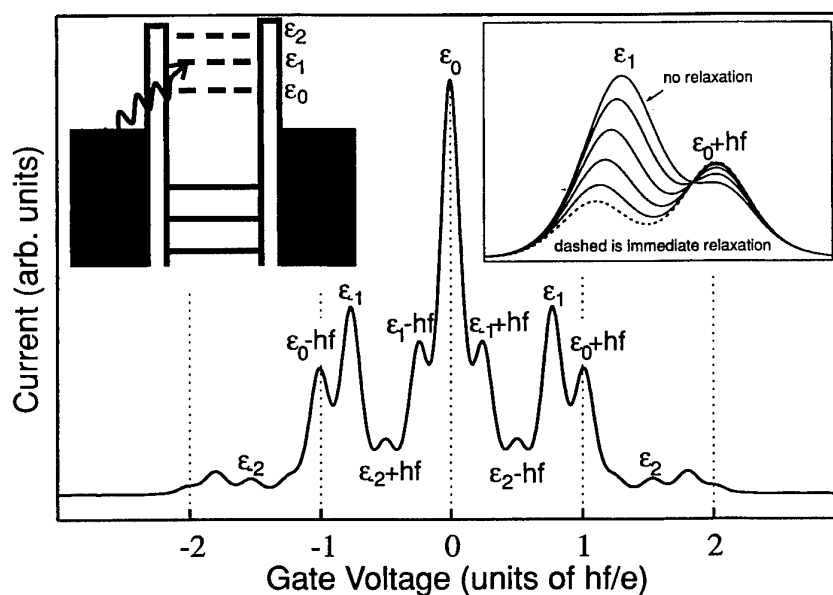


Figure 2. Expansion for the curve $\alpha = 1$ from figure 1. The right inset shows the effects of increasing relaxation. The relaxation rates divided by the tunnel rate are 0, 0.1, 0.35, 1, 3.5 and infinite. The left inset sketches a photon absorption process to overcome the Coulomb gap.

The d.c. current I was measured as a function of the d.c. gate voltage V_g in the presence of a microwave signal. The data shown in figure 3 were taken at 45 GHz for different microwave powers corresponding to α ranging from 1 to 3. The photon energy $\hbar f = 185 \mu\text{eV}$ is indicated by the vertical dotted lines in the figure. Bumps with a spacing corresponding to the photon energy are clearly seen. We interpret these bumps as photon side peaks. Additional structure due to excited OD states in between photon side peaks is washed out, probably because the difference between the photon energy and the OD

splitting is comparable to temperature. At higher powers the neighbouring Coulomb peak starts to lift the current at the right of the peak.

The broken curves are fits. The lower curve has $\alpha = 1$ and no relaxation. The upper curve fitted best with $\alpha = 1.4$ and 50% relaxation. However, it is not clear whether the incorporation of relaxation is really required to explain the data.

Our measurements do show for the first time PAT through discrete OD states. With a lower electron temperature the side peaks should become more pronounced

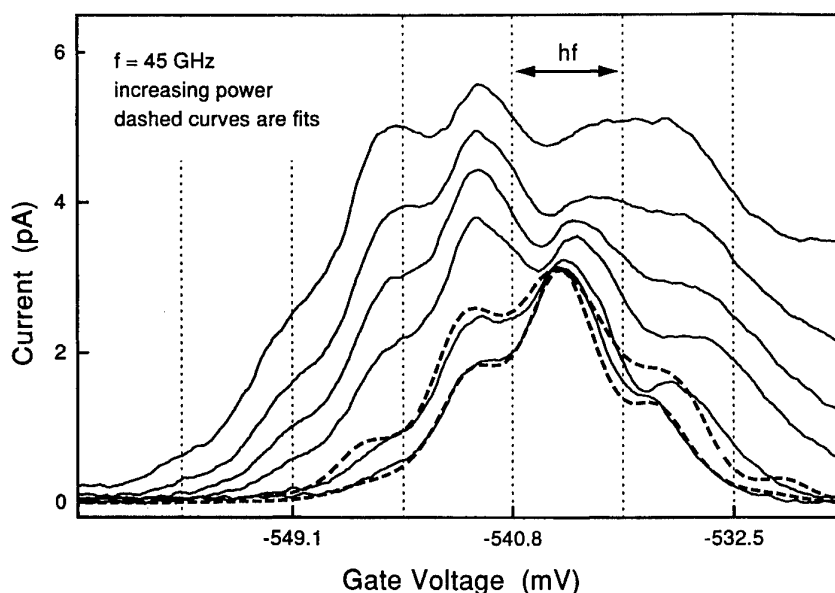


Figure 3. A measurement of a single Coulomb peak in the presence of microwaves with $f = 45$ GHz. The different curves are for increasing power. The two broken curves are fits for $\alpha = 1$ and 1.4 .

and a more detailed comparison between simulations and experiments may give information concerning relaxation processes. Future work will also include PAT in double dots with zero-dimensional states. Here it should be possible to attain a higher resolution [10]. Double dots are also promising for studying coherent tunnel processes involving PAT [11].

Acknowledgments

We would like to thank P Lukey, S F Godijn and J E Mooij for discussions and experimental help. Also we thank Philips Laboratories and C T Foxon for providing the heterostructures. LPK was supported by the Royal Netherlands Academy of Arts and Sciences.

References

- [1] Tien P K and Gordon J R 1963 *Phys. Rev.* **129** 647
- [2] Patel N K *et al* 1991 *Proc. 20th Int. Conf. on the Physics of Semiconductors* ed E M Anastassakis and J D Joannopoulos (Singapore: Scientific) p 2371
- [3] Wyss R A *et al* 1993 *Appl. Phys. Lett.* **63** 1522; 1995 *Appl. Phys. Lett.* **66** 1144
- [4] Jansen T J B M *et al* 1994 *J. Phys.: Condens. Matter* **6** L163
- [5] Guimarães P S S *et al* 1993 *Phys. Rev. Lett.* **70** 3792
- [6] Keay B J *et al* 1995 *Phys. Rev. Lett.* **75** 4098; 1995 *Phys. Rev. Lett.* **75** 4102
- [7] Kouwenhoven L P *et al* 1994 *Phys. Rev. B* **50** 2019; 1994 *Phys. Rev. Lett.* **73** 3443
- [8] Blick R H *et al* 1995 *Appl. Phys. Lett.* **67** 3924
- [9] Mixing is described by J Iñarrea, G Platero and C Tejedor 1994 *Phys. Rev. B* **50** 4581
- [10] Averin D V *et al* 1991 *Phys. Rev. B* **44** 6199
- [11] Beenakker C W J 1991 *Phys. Rev. B* **44** 1646
- [12] Wan J C *et al* 1991 *Phys. Rev. B* **43** 9381
- [13] Bruder C and Schoeller H 1994 *Phys. Rev. Lett.* **72** 1076
- [14] van der Vaart N C *et al* 1995 *Phys. Rev. Lett.* **74** 4702
- [15] Stoof T H and Nazarov Yu V 1995 *Phys. Rev. B* **51**
- [16] Stafford C A and Wingreen N S 1996 *Phys. Rev. Lett.* **76** 1916

Electronic structure and optical properties of self-assembled quantum dots

Pawel Hawrylak and Arkadiusz Wojs†

Institute for Microstructural Sciences, National Research Council of Canada,
Ottawa, Canada K1A 0R6

Abstract. The electronic structure of lens-shaped InGaAs self-assembled quantum dots (SADs) is studied as a function of the dot size, the confining potential and the magnetic field. Numerical calculations show the formation of the electronic shell structure to be well approximated by a Fock–Darwin energy spectrum. Many-body effects in electronic and optical properties of SADs charged with electrons and/or excitons are investigated using exact diagonalization techniques. The capacitance, infrared and interband absorption/emission spectra are calculated as a function of the number of particles, the size and the magnetic field.

1. Introduction

Self-assembled dots (SAD) [1] are small quasi-two-dimensional semiconductor structures. Their shape and hence their electronic properties depend on growth conditions. Dots in the shape of pyramids [2, 3], disks [4] and lenses [1, 5–7] have been reported, although the actual determination of the shape is not definite. The SAD described here are the lens-shaped structures spontaneously formed on a narrow InGaAs wetting layer (WL) and surrounded by the GaAs barriers [1, 5–7]. We study the dependence of energy levels of SAD on size, depth of confining potential and the magnetic field in the effective mass approximation. In our SAD both electrons and valence band holes are confined and interband magneto-optics is a useful tool in relating structural parameters to the photoluminescence (PL) and absorption spectra [6, 7]. For many applications, the SAD must be charged with either many electrons or many excitons. We investigate the effects of electron–electron and electron–hole interactions on electronic and optical properties of SAD containing many electrons and/or excitons [2, 7–9] using exact diagonalization techniques. The single-electron capacitance (SECS), far-infrared (FIR) and PL/absorption spectra are calculated as a function of the number of particles, the size and the magnetic field. Results are compared with SECS and FIR experiments by Drexler *et al* [5] and PL experiments by Raymond *et al* [7].

2. Single-particle states

A schematic picture of a lens-shaped $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ SAD formed on a wetting layer (WL) of thickness t_w , and

modelled as a part of a sphere with fixed height h and radius at the base s is shown in figure 1. SAD with different sizes are characterized by a constant ratio h/s [6, 7]. The carriers, confined to a narrow WL quantum well, are further localized in the area of the dot due to the effectively increased thickness of the layer. The parameters of WL and SAD material enter through the effective Rydberg constant $Ry^* = m_e e^4 / 2\epsilon^2 \hbar^2$ and the effective Bohr radius $a_B^* = \epsilon \hbar^2 / m_e e^2$, with m_e and ϵ being the electron effective mass and the dielectric constant respectively. These parameters include all effects due to strain, discontinuities in the effective mass, dielectric constants, etc. The effects of strain are approximated by calculating a uniform hydrostatic-pressure shift and a uniaxial stress-induced valence-band splitting. Assuming a conduction-band offset of 67%, the resulting electron confining potential is $V_0 = 350$ meV, the effective mass of strained $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ $m_e = 0.067$ and the dielectric constant $\epsilon = 12.5$.

Two approximate methods of solving the three-dimensional (3D) Schrödinger equation were used [10]: (a) full 3D numerical diagonalization for SAD embedded in a large infinite-wall disk and (b) an adiabatic approximation.

In the physically intuitive adiabatic approximation, the 3D problem is reduced to an effective two-dimensional (2D) problem since the electron's wavefunction is strongly confined in the narrow quantum well of the WL. With coordinates r in the plane and z perpendicular to the WL, the wavefunction is written as $\psi(r, \theta, z) = \frac{1}{\sqrt{2\pi}} e^{im\theta} g_r(z) f_m(r)$. Here $g_r(z)$ is a slowly varying function of r in a confining potential $V_e(r, z)$. In each angular momentum channel m the wavefunctions g and f_m satisfy a set of equations:

$$\left(-\frac{\partial^2}{\partial z^2} + V_e(r, z) \right) g_r(z) = E_0(r) g_r(z) \quad (1)$$

† Permanent address: Institute of Physics, Technical University of Wrocław, 50-370 Wrocław, Poland.

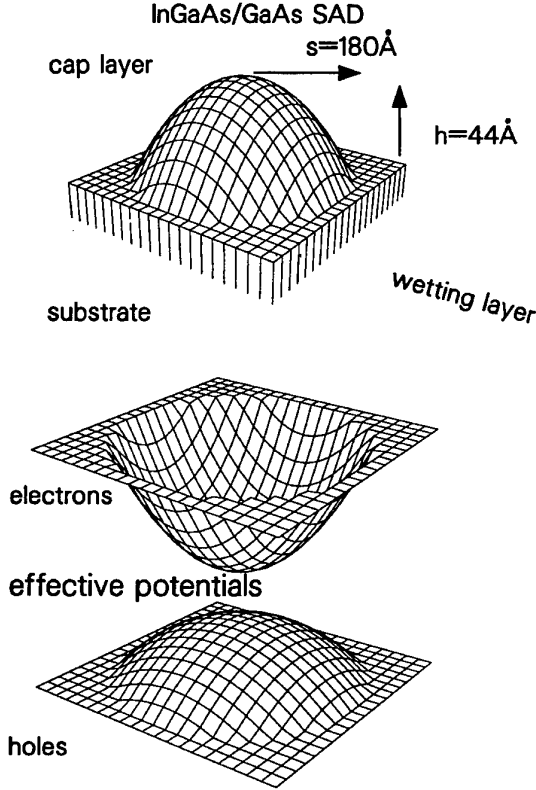


Figure 1. Top: a schematic picture of the InGaAs/GaAs SAD modelled as a part of a sphere on a narrow WL. Bottom: effective lateral confining potential for electrons and for holes.

$$\left[-\frac{1}{r^2} \left(r \frac{\partial}{\partial r} r \frac{\partial}{\partial r} - m^2 \right) + E_0(r) \right] f_m(r) = E f_m(r). \quad (2)$$

We first find the energy $E_0(r)$ corresponding to the motion in the z direction at the distance r from the centre of the dot, shown as the effective 2D potential in figure 1. The radial motion in the effective potential $E_0(r)$ is next solved exactly for each angular momentum channel for both *bound* and *scattered* states.

In figure 2 we show the dependence of the energy spectra of bound states on the radius of the dot s , with a fixed ratio $h/s = 0.24$ and $V_0 = 350$ meV. As the size of the dot exceeds $s \approx 80$ Å, bound states with significant binding energy become visible. The states tend to bunch into groups, forming well separated shells. The number of bound states and the spacing of energy levels compare well with [5, 7]. The spacing of energy levels and the wavefunctions are very well fitted by Fock–Darwin (FD) energy levels, as discussed below.

The FD states $|nm\rangle = (a^\dagger)^n (b^\dagger)^m |00\rangle / \sqrt{n!m!}$ are those of a pair of harmonic oscillators with energies $\Omega_\pm = \frac{1}{2}(\Omega \pm \omega_c)$, where $\Omega^2 = \omega_0^2 + 4\omega_c^2$, $\omega_c = eB/cm_e$ is the cyclotron energy, and ω_0 measures the effective confinement energy. Associating index n with frequency Ω_+ and index m with frequency Ω_- , the energy E_{nm}^e and orbital angular momentum R_{nm} of the state $|nm\rangle$ are $R_{nm} = m - n$ and $E_{nm}^e = \Omega_+(n + \frac{1}{2}) + \Omega_-(m + \frac{1}{2})$. The Zeeman energy is very small and can be neglected.

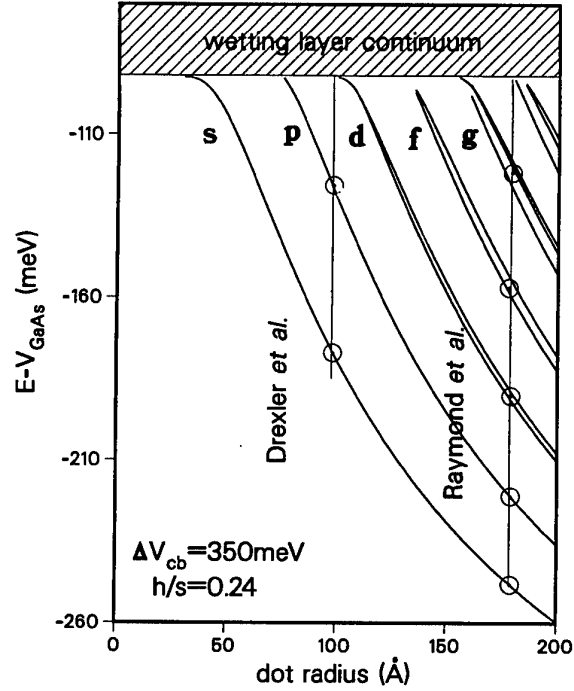


Figure 2. Dependence of the electronic energy spectrum of the SAD on its size.

The eigenstates are doubly degenerate due to spin σ . The valence band holes are treated in the effective mass approximation as a positively charged particle with angular momentum $R_{mn}^h = n - m$, opposite to the electron, and FD energies $E_{mn}^h = \Omega_+^h(n + \frac{1}{2}) + \Omega_-^h(m + \frac{1}{2})$ (ignoring the semiconductor gap E_G). For special values of the magnetic field B_p such that $\Omega_+ = p\Omega_-$, the energy spectrum of electrons (holes) $E_{mn} = \Omega_-(n + pm + 1)$ is degenerate for $n + pm = t$, with t labelling electron (hole) shells and g_t the degeneracy of each shell. Level crossing takes place for intermediate magnetic fields. For $B = 0$, the degenerate shells can be labelled as s, p, d, ... In a strong magnetic field, degeneracies are removed and electron (hole) energies $E(m)$ increase linearly with angular momentum $R = m$.

3. Many-particle states

With a composite index $j = [m, n, \sigma]$ describing the FD states, the Hamiltonian of the interacting electron (electron-hole) system may be written in a compact form:

$$\begin{aligned} H = & \sum_i E_i^e c_i^\dagger c_i + \sum_i E_i^h h_i^\dagger h_i - \sum_{ijkl} \langle ij | v_{eh} | kl \rangle c_i^\dagger h_j^\dagger h_k c_l \\ & + \frac{1}{2} \sum_{ijkl} \langle ij | v_{ee} | kl \rangle c_i^\dagger c_j^\dagger c_k c_l \\ & + \frac{1}{2} \sum_{ijkl} \langle ij | v_{hh} | kl \rangle h_i^\dagger h_j^\dagger h_k h_l. \end{aligned} \quad (3)$$

The operators $c_i^\dagger (c_i)$, $h_i^\dagger (h_i)$ create (annihilate) the electron or valence band hole in the state $|i\rangle$ with the single-particle energy E_i . The two-body Coulomb matrix elements are

$\langle ij|v|kl\rangle$ for electron–electron (ee), hole–hole (hh) and electron–hole (eh) scattering respectively [11, 12].

The eigenstates $|\nu\rangle$ of the electron (electron–hole) system with N electrons (excitons) are expanded in products of the electron and hole configurations $|\nu\rangle = (\prod_{j=1}^N c_j^+) (\prod_{k=1}^N h_k^+) |\nu\rangle$. Up to 30 single-particle states per dot, including spin, are used in calculations. The configurations are labelled by total angular momentum R_{tot} and the z component of total spin S_z^{tot} . For electrons, numerical diagonalization of up to $N = 6$ electrons was carried out for Hilbert spaces with different total angular momentum. For excitons, we concentrate on the optically active subspace of $R_{tot} = 0$ and $S_z^{tot} = 0$. A combination of exact diagonalization techniques (for up to $N = 6$ excitons) in configuration space coupled with the Hartree–Fock approximation extended calculations up to $N = 20$ excitons.

4. Capacitance and infrared spectroscopy

Drexler *et al* [5] reported SECS and FIR absorption measurements of charged self-assembled $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ SAD in a magnetic field. The dots were charged with electrons filling the s and p electronic shells and infrared (IR) spectroscopy was used to study the electronic excitations of the dots as a function of the number of electrons and the magnetic field.

SECS measures the chemical potential $\mu(N)$ of the dot [13, 14]. The chemical potential $\mu(N) = GSE(N) - GSE(N-1)$ is the difference in the ground-state energy (GSE) of the dot due to the addition of a single electron. In figure 3(a) we show the calculated SECS spectra corresponding to charging of the p shell of SAD with increasing magnetic field.

For $N < 4$ electrons the large quantization of kinetic energy prevents ‘magic state’ transitions [13] and the ground state (GS) is simply equivalent to the minimum kinetic energy configuration. The effects of electron–electron interactions and the magnetic field begin to play a role for $N = 4$ electrons, i.e. two electrons in a partially filled p shell. For low magnetic fields the two ‘core electrons’ of the s shell are frozen in a spin singlet state while the two electrons on the p shell occupy degenerate FD orbitals $|01\rangle$ and $|10\rangle$. The GS is a spin triplet, zero total angular momentum state $R = 0, S = 1$. The triplet state lowers its energy by an exchange-interaction term $\langle 01; 10|V_{ee}|01; 10\rangle$ while the spin singlet state $R = 0, S = 0$ always increases its energy by the same amount. With increasing magnetic field the triplet state $R = 0, S = 1$ begins to compete with a finite angular momentum spin singlet state $R = 2, S = 0$. In the $R = 2$ state both p electrons occupy the lower energy orbital $|10\rangle$. At $B \approx 2.8$ T the gain in exchange energy of the triplet configuration $R = 0, S = 1$ is overtaken by an increase of kinetic energy and the system makes a transition into a spin-singlet lower kinetic energy configuration. The transition is marked in figure 3 with an arrow. Contrary to the expectation for noninteracting electrons, the addition energy for a triplet *increases* with increasing magnetic field up to $B = 2.8$ T. For $N = 5, 6$ the changes of slope

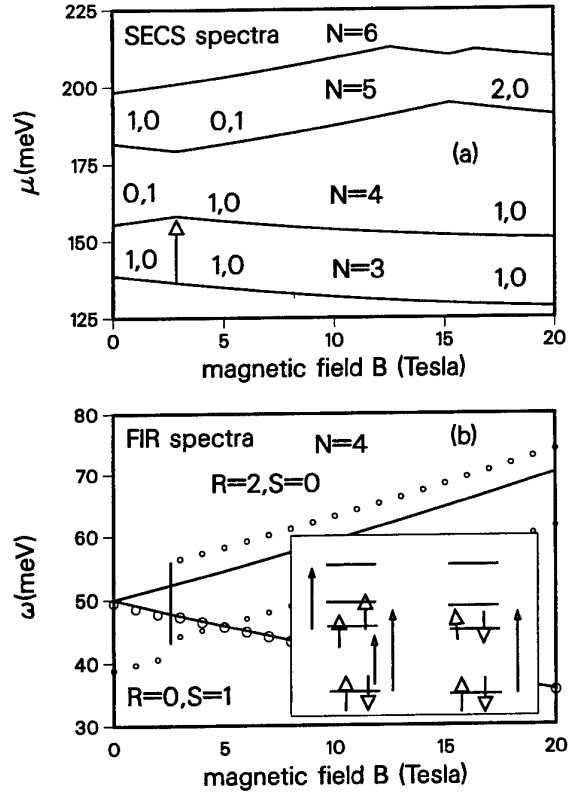


Figure 3. (a) SECS spectra of a p shell of a SAD. (b) FIR spectra of a SAD with $N = 4$ electrons. The inset shows single-particle configurations before and after transition and arrows indicate excitations.

at higher magnetic fields are due to the crossing of FD levels $|0, 1\rangle$ and $|2, 0\rangle$. The splitting of SECS spectra of p states by the magnetic field has been observed in [5] but the consequences of electron–electron interactions discussed here had not yet been observed.

The excited states of SAD also reflect the electronic structure and the number of electrons in the dot. For a parabolic confinement only the CM excitations with frequencies Ω_+ and Ω_- (generalized Kohn’s theorem) [15, 18] can be measured in the FIR. In SAD, a finite number of confined FD levels leads to additional transitions in the IR spectrum related to the magnetic-field-induced changes in the GS.

The FIR absorption for N electrons can be conveniently expressed in terms of the FD ladder operators a and b [12, 17]:

$$A(\omega) \propto \sum_f |\langle f | \sum_{j=1}^N (a_j + a_j^\dagger + b_j + b_j^\dagger) | i \rangle|^2 \delta(E_f - E_i - \omega) \quad (4)$$

where $|i\rangle$ is the initial (ground) state and the summation is over all bound final states $|f\rangle$. IR radiation connects only the states with the same S_z^{tot} and total angular momenta different by ± 1 . We have shown in figure 3(b) the magnetic field evolutions of the IR spectra calculated for the SAD with $N = 4$ electrons. The area of each circle is proportional to the intensity $A(\omega)$. The solid lines show

the transition energies Ω_{\pm} of the noninteracting system and a vertical line marks the transition in the GS. The GS single-particle configurations before and after the transition at $B = 2.8$ T, together with low-energy ($\approx \Omega_-$) and high-energy ($\approx \Omega_+$) excitations, are indicated in the inset. The FIR spectrum reflects the transition in the GS of the dot in terms of an additional transition. In experiment, the energy $\Omega_+ \approx \omega_c$ will probably be visible due to excitations to scattered states in Landau levels in the WL.

5. Many excitons in highly excited quantum dots

We now turn to the effects of exciton–exciton interaction on optical properties of a highly excited SAD [7]. This SAD has 15 bound states (shells s–g) and can be filled with up to 30 electrons and holes.

The calculations for up to $N = 6$ electron–hole pairs were carried out exactly and a combination of numerical diagonalization in a partially filled shell and the Hartree–Fock approximation extended calculations up to $N = 20$ excitons. Due to the large confinement, the lowest kinetic energy configurations are an excellent approximation in the case of filled shells. When electrons and holes partially fill up a degenerate shell, the states and energies are completely determined by their mutual interactions and exact numerical calculations are necessary.

In figure 4 we show an example of the effect of population of electrons and holes on the absorption (addition) spectrum. The dotted line follows the chemical potential (addition spectrum) of excitons. The single-exciton spectrum consists of peaks shifted from transitions in the noninteracting system by electron–hole interactions. The two-exciton spectrum is richer and shows the bi-exciton binding of ≈ 5 meV.

When excitons are added into the p shell, the energy to add excitons is almost constant. The same is true for d and f shells as indicated by the chemical potential. The actual positions of the p, d, f shells are, however, shifted from corresponding peaks in the noninteracting and single-exciton spectrum. The shift is equivalent to bandgap renormalization in quantum dots. The remarkable steps in the addition spectrum of excitons indicate that excitons form a weakly interacting gas of bi-excitons and excitons.

For most quantum dots where electrons and holes are confined in the same physical area, the electron and hole interactions are very symmetrical. For example, in the sample calculated here $v_{ee}/v_{eh} = v_{eh}/v_{hh} = 1.04$. For almost symmetrical interactions, when the Hamiltonian is restricted to a single degenerate shell t , the commutator of the Hamiltonian and the interband polarization operator $P^+ = \sum_i c_{i\uparrow}^\dagger h_{i\downarrow}^\dagger$ can be approximated as $[H_t, P^+] \approx E_X^t P^+$, where $E_X^t = E_t^e + E_t^h - \sum_{jk} g_t^{-1} \langle jj | v_{eh} | kk \rangle$ is an approximate exciton binding energy. The quantum number j ($-t \leq j \leq t$) denotes the angular momentum on a given shell. This commutation relation is a manifestation of hidden symmetry [16]. One can construct coherent N exciton states $(P^+)^N |v\rangle$ as eigenstates of P^2 . Due to hidden symmetry these states are also eigenstates of the shell Hamiltonian with energies $E(N_\sigma) = N_\sigma E_X^t$. The energy of these states is just the sum of energies of noninteracting

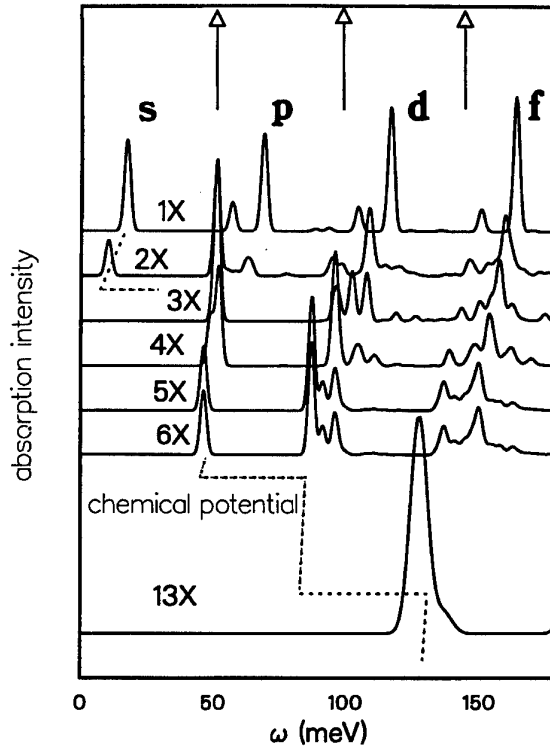


Figure 4. Absorption spectrum for SAD with different number of excitons. The broken line shows the chemical potential of excitons. Arrows indicate transitions in the noninteracting system.

excitons. The coherent states turn out to be excellent approximations to exact GSs with corresponding overlaps of 100% for shells s and p, and 99.8% and 99.2% for shells d and f. For realistic calculations including shell–shell scattering we find that excitons form a gas of weakly interacting bi-excitons and excitons and the steps in the addition spectrum survive, as shown in figure 4.

6. Conclusions

To summarize, the calculated single-particle states of lens-shaped SAD can be well approximated by FD levels. This simplification allowed us to carry out calculations of many-body effects in SAD charged with electrons and with excitons. The calculated addition SECS and FIR spectra in a magnetic field reflect both the single-particle energy levels and the effects of electron–electron interactions. The addition spectrum of excitons shows plateaus reflecting (a) the shell structure of single-particle energy levels, (b) hidden many-particle symmetries and (c) bandgap renormalization.

Acknowledgments

Discussions with S Fafard, S Raymond, S Charbonneau, P Poole, G C Aers and J P Kotthaus are gratefully acknowledged.

References

- [1] Petroff P M and Denbaars S P 1994 *Superlatt. Microstruct.* **15** 15; for a recent review see *Proc. Int. Conf. on Modulated Semiconductor Structures (Madrid, 1995)*
- [2] Grundmann M *et al* 1995 *Phys. Status Solidi* **188** 249
- [3] Marzin J Y and Bastard G 1994 *Solid State Commun.* **92** 437
- [4] Notzel R *et al* 1995 *Appl. Phys. Lett.* **66** 2525
- [5] Drexel H *et al* 1994 *Phys. Rev. Lett.* **73** 2252
- [6] Fafard S *et al* 1994 *Appl. Phys. Lett.* **65** 1388
- Leon R *et al* 1995 *Appl. Phys. Lett.* **67** 521
- [7] Raymond S *et al* to be published
- Fafard S *et al* 1995 *Phys. Rev. B* **52** 5752
- [8] Bockelmann U *et al* 1996 *Phys. Rev. Lett.* **76** 3622
- [9] Bayer M *et al* 1995 *Phys. Rev. Lett.* **74** 3439
- [10] Wojs A, Hawrylak P, Fafard S and Jacak L 1996 *Phys. Rev. B* **54** 15
- [11] Wojs A and Hawrylak P 1995 *Phys. Rev. B* **51** 880
- [12] Hawrylak P 1993 *Solid State Commun.* **88** 475
- [13] Hawrylak P 1993 *Phys. Rev. Lett.* **71** 3347
- [14] Ashoori R C *et al* 1993 *Phys. Rev. Lett.* **71** 613
- [15] For recent reviews and references see Kastner M 1993 *Phys. Today* **24**
- Chakraborty T 1992 *Comment. Condens. Matter Phys.* **16** 35
- [16] Lerner I V and Lozovik Yu E 1981 *Zh. Eksp. Teor. Fiz.* **80** 1488 (Engl. Transl. 1981 *Sov. Phys.-JETP* **53** 763)
- Paquet D, Rice T M and Ueda K 1985 *Phys. Rev. B* **32** 5208
- MacDonald A H and Rezayi E H 1990 *Phys. Rev. B* **42** 3224
- Bychkov Yu A and Rashba E I 1991 *Phys. Rev. B* **44** 6212
- [17] Pfannkuche D, Gudmundsson V, Hawrylak P and Gerhards R R 1994 *Solid-State Electron.* **37** 1221
- [18] Kohn W 1961 *Phys. Rev.* **123** 1242
- Brey L, Johnson N and Halperin B 1989 *Phys. Rev. B* **40** 647
- Maksym P and Chakraborty T 1990 *Phys. Rev. Lett.* **65** 108

Growth and characterization of self-assembled Ge-rich islands on Si

G Abstreiter[†], P Schittenhelm[†], C Engel[†], E Silveira[†],
A Zrenner[†], D Meertens[‡] and W Jäger[‡]

[†] Walter Schottky Institut, Techn. Univ. München, Am Coulombwall,
D-85748 Garching, Germany

[‡] Institut für Festkörperforschung, Forschungszentrum Jülich GmbH,
D-52452 Jülich, Germany

Abstract. Ge-rich islands have been grown on Si (100) substrates by molecular beam epitaxy. Their density and size distribution are analysed by atomic force microscopy as a function of growth temperature, growth rate and Ge coverage. Overgrown islands have been studied by transmission electron microscopy, Raman scattering and photoluminescence. The first results of photocurrent spectroscopy on Si/Ge/Si pin diodes show the expected shift of the energy gap. Based on these results, novel device applications of Ge-rich islands in Si are proposed.

1. Introduction

Self-assembled semiconductor islands which are embedded coherently in a different semiconductor host material have attracted considerable interest in recent years. Such islands are achieved by epitaxial growth of lattice mismatched systems which show a Stranski–Krastanov-type of growth mode under certain conditions. This means that three-dimensional islands are formed after deposition of typically a few monolayers (ML) on a substrate with a different lattice constant which leads to a reduction of the strain energy. Their size distribution is rather narrow near the onset of island formation and there is some hope that this approach may lead to the achievement of a homogeneous ensemble of quantum dots where the discrete atomic-like density of states can be used for quantum device applications. The most widely studied system so far is (InGa)As/(AlGa)As for which various quantum effects have already been demonstrated [1–6]. Other III–V semiconductor systems studied so far involve InP/InGaP [7, 8], AlInAs/GaAlAs [9] and GaSb/GaAs [10]. In this contribution we discuss recent results which have been achieved for self-assembled Ge-rich islands grown by molecular beam epitaxy (MBE) on Si (100) substrates. This system was studied earlier with various growth methods and characterization techniques in different regimes of island formation [11–16].

The paper is organized as follows: the sample structures and the experimental set-up are described in section 2. Results of the structural analysis of the islands are presented in section 3. In section 4 we show the first results obtained with pin diode structures; finally we discuss some concepts for future device applications based on such islands in section 5.

2. Experimental set-up

All samples were grown on Si (100) substrates by solid source MBE in a commercial Riber Siva 32 system. Details of substrate preparation and growth have been reported elsewhere [16, 17]. For analysis with atomic force microscopy (AFM) the samples typically consist of a 150 nm Si buffer with 4 to 12 ML Ge on top. Samples for photoluminescence (PL) spectroscopy were grown with 90 nm Si cap layers. For photocurrent (PC) spectroscopy, a few periods, each consisting of a 40 nm Si spacer and 12 ML Ge, were grown on a 15 m Ω cm boron-doped Si substrate and capped by a 150 nm n-doped (1×10^{19} cm $^{-3}$ P) top contact. PL measurements were performed at 4 K, using an Ar $^{+}$ laser for excitation. The signals were analysed with a single-grating monochromator together with a liquid-nitrogen-cooled Ge detector using a standard lock-in technique. For PC spectroscopy, a tunable broadband light source was used for excitation. The photocurrent was measured using a current amplifier together with the lock-in technique. AFM images were recorded in contact mode with a TopoMetrix AFM in ambient air. Transmission electron microscopy (TEM) was performed on [110] cross sections of samples prepared by mechanical grinding and subsequent ion milling.

3. Structural analysis

In order to get information on the influence of the growth parameters on the formation and the structural properties of self-organized Ge islands on Si, a series of samples with different Ge coverages, different Ge concentrations, and varying growth temperatures and growth rates have been prepared. As an example we show two AFM images in figure 1. The sample with a high density of islands consists

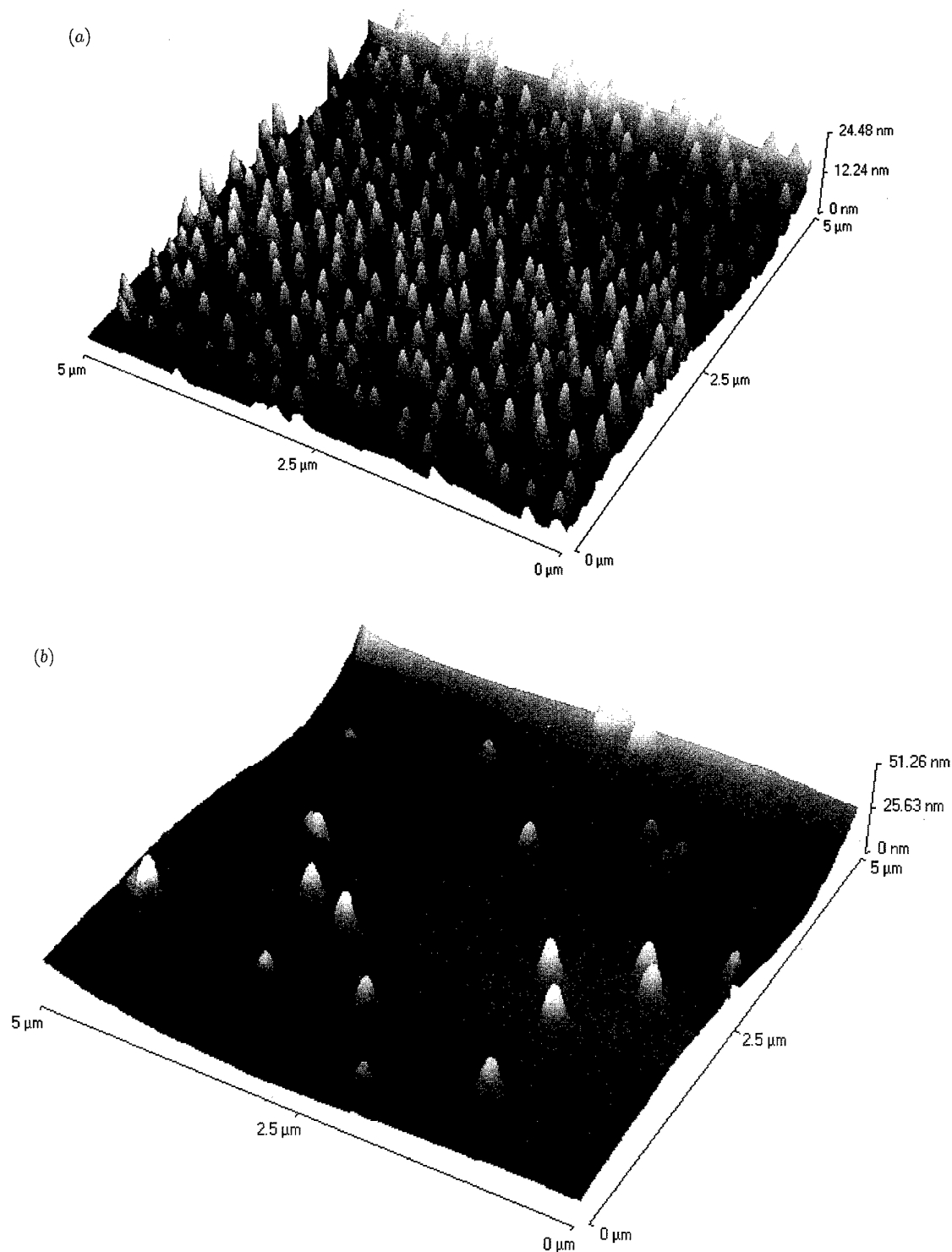


Figure 1. AFM images of Ge islands on Si. The top part shows islands formed from nominally 6 ML of Ge grown at 670°C, the lower part 5 ML grown at 740°C.

of nominally 6 ML of Ge grown at a substrate temperature of 670°C. Under such conditions the typical island size is about 15 nm in height and 150 nm in diameter. However, the diameter and the height of the islands is significantly overestimated by the AFM measurements as the extracted

values are a convolution of the islands and the tip, the radius of which is 50 nm and thus roughly that of the islands. Furthermore, oxidation of the islands in ambient air also increases their volume. The island density is approximately 1.5×10^9 islands/cm². The size and density of these islands

are strongly influenced by the coverage and the growth conditions. The lower part of figure 1 shows an example for a nominal Ge coverage of 5 ML at a growth temperature of 740°C. The island density is strongly reduced and the sizes are larger due to the higher growth temperature. More generally we observe the following behaviour: first Ge islands are formed at a coverage of about 5 ML at growth temperatures above 560°C. The density increases strongly when more Ge is deposited concomitant with a more inhomogeneous size distribution. The average island size also increases with growth temperature or with a reduction in growth rate. Some such dependences are shown in figure 2, where the areal density and the average island diameter is plotted versus growth temperature and growth rate for a nominal layer thickness of 6 ML as determined from AFM images. Increasing growth temperature as well as a decrease in the Ge flux leads to an enhanced diffusion length of the Ge adatoms on the sample surface. As the formation of islands allows the system to reduce its total energy by elastic relaxation of strain in the islands [11], all adatoms will preferably be incorporated in existing islands. Therefore, with increasing diffusion length of the Ge adatoms more of them reach the energetically favourable islands, which leads to a reduction of island density concomitant with an increase in their size. This is clearly observed both with increasing temperature (figure 2(a)) and for a reduced growth rate, as shown in figure 2(b), for a fixed growth temperature of 740°C. Simultaneously, the average diameter of the islands increases. These results are in good qualitative agreement with those obtained for InGaAs islands on GaAs [18, 19].

For fixed growth conditions the size and density of islands also depend strongly on the Ge coverage and/or on the Ge content in the surface layer. This is shown for example in figure 3, where the island height is plotted versus layer thickness between 6 and 18 ML of 85% Ge and for a series of samples for which the nominal layer thickness was kept constant at 12 ML but the Ge content was varied between 50% and 100%. In both cases the island size increases with the average amount of Ge in the film. The areal density remains nearly unchanged when the Ge content is altered for 12 ML thickness and fixed growth parameters. The island height, however, increases continuously with increasing Ge content, as well as island diameter. Only the lattice mismatch and consequently the strain is changed in this series of samples. These results therefore suggest that the shape of the islands can be influenced by the strain in the system. This has not yet been analysed in detail, but it has been shown, however, that the critical thickness for island formation increases with reduced strain. This will be discussed in more detail later in connection with PL results.

In order to be able to study the optical and electrical properties of such islands, it is necessary to overgrow them with Si cap layers. This is expected to change the shape and composition especially when the overgrowth is performed at high temperatures. Segregation and interdiffusion may play an important role in the final Si/Ge/Si island structures. To get some information on these effects we performed both TEM analysis and Raman experiments on selected

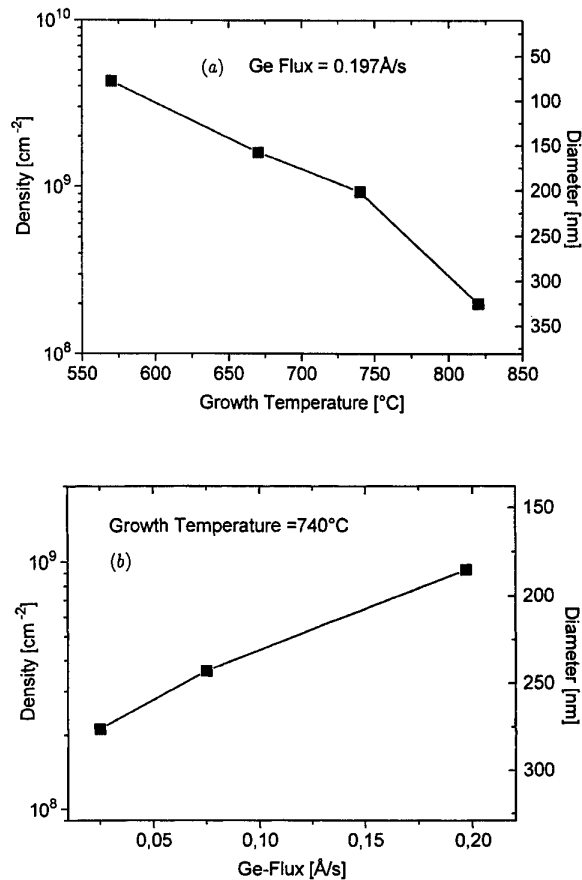


Figure 2. Areal density and average diameter of Ge islands for 6 ML coverage, depending on growth temperature (a) and Ge flux (b).

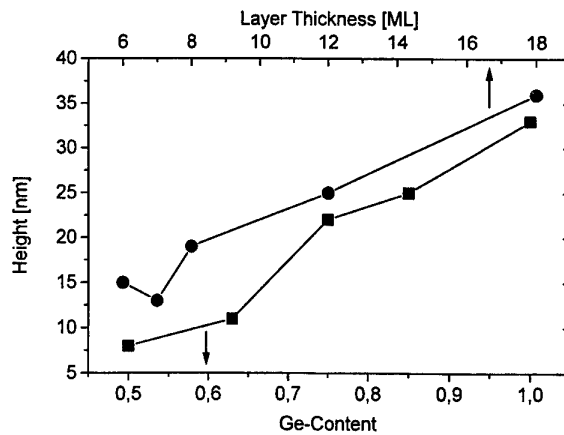


Figure 3. Height of islands versus Ge content for a coverage of 12 ML (squares) and versus coverage for a nominal Ge content of 85% (circles).

samples. Cross-sectional TEM imaging reveals that no misfit dislocations are introduced by the overgrowth of Ge islands with Si for Ge thicknesses of about 6 ML and growth temperatures of 740°C. An example is shown in figure 4. This is not valid any more for islands formed from 12 ML Ge, where some defects or misfit dislocations

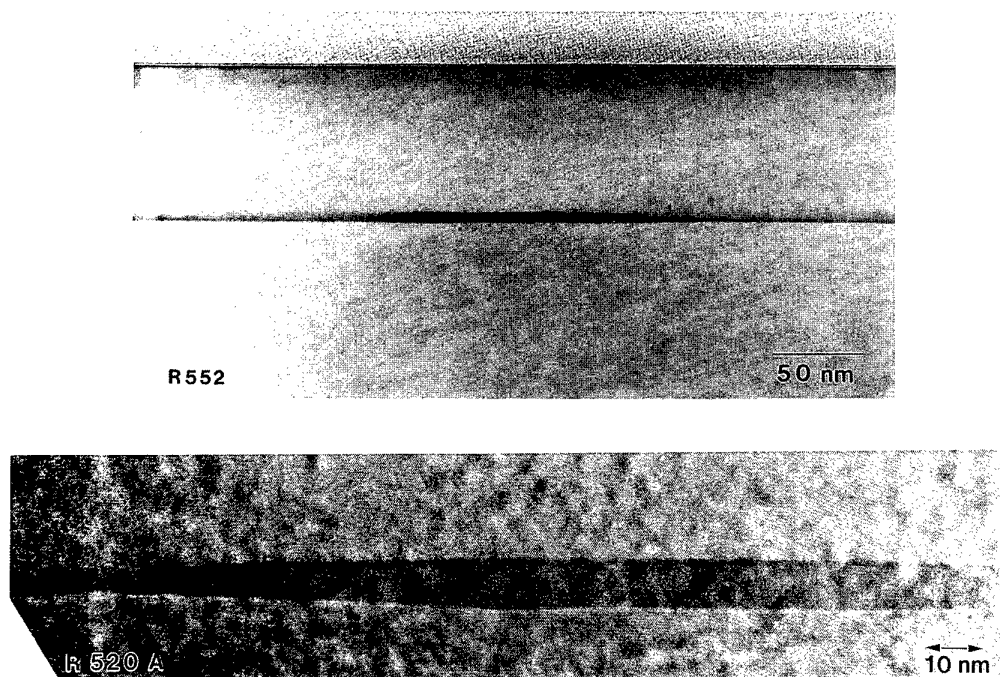


Figure 4. Selected examples of cross-sectional TEM images of Ge islands in Si. The top part shows an island of a nominally 6 ML thick Ge layer with lower magnification. The wetting layer on the left- and right-hand sides of the island has a sharp lower and a washed-out upper interface. In the lower image, part of an island formed of a 12 ML Ge layer is shown with higher magnification. Some faceting is visible at the edge of the island, also into the substrate. The diameter of the island is 250 nm.

have been detected on top of some islands. The actual shapes of the embedded islands are changed in comparison with the AFM results for all thicknesses. While the lateral dimensions are roughly comparable, the height seems to be reduced by about a factor of two after overgrowth. A similar effect has already been reported for InGaAs islands in GaAs [20], which is not yet fully understood but which is probably related to segregation and diffusion. One has to keep in mind, however, that in AFM in ambient air one measures the convolution of the tip shape with an oxidized free surface. That interdiffusion with the neighbouring Si layers indeed plays an important role, especially at such high growth temperatures, is also evident from some faceting of the islands into the underlying Si as observed in high-resolution TEM (see figure 4(b)). Also the upper interface of the wetting layer is smeared out due to Ge segregation during overgrowth.

The observation of these effects raises the important question about the actual Ge content of the embedded islands. We have used Raman measurements and PL to get some information on their composition. For samples grown at $T = 740^\circ\text{C}$ and for nominally 12 ML Ge we are able to detect optical phonon modes due to both Si-Ge and Ge-Ge vibrations, indicating that the nominally pure Ge islands contain a considerable amount of Si. The Raman measurements average, however, over the interior of the islands, the interfaces and the wetting layer in between the islands, and it is thus difficult to extract quantitative information on the actual composition of the dots. The energetic positions and the intensity ratios are best fit with

roughly the same number of Ge-Ge and Si-Ge bonds in the sample, suggesting an average composition of about 50% in the whole layer. The mentioned averaging probably overestimates the Si content in Ge.

Another powerful technique for studying the properties of overgrown islands is PL. Figure 5 shows a series of PL spectra for samples grown at 740°C with an increasing Ge layer thickness. The pair of lines at higher energies, labelled 2D QW, are the no-phonon (NP) line and its TO-phonon replica of the narrow Ge-rich quantum well. The broader doublet observed at energies below 0.95 eV, labelled 3D islands, is attributed to electron-hole recombination inside the Ge-rich islands or at their interface to the surrounding Si. The PL peak energies are plotted versus nominal Ge layer thickness in figure 6. At low coverages, the PL signals are red-shifted with increasing thickness, reflecting the decreasing quantum confinement in the 2D wetting layer. A further increase of the Ge coverage leads to a blue-shift of the PL of the 2D quantum well. This was attributed to an increasing quantum confinement due to lateral diffusion of Ge towards the dots, resulting in a thinning of the underlying 2D wetting layer [16]. The onset of this blue-shift can be regarded as a critical thickness for the formation of islands. Figure 7 shows this critical thickness in dependence on the Ge content. The expected increase for lower Ge contents, i.e. less lattice mismatch, is obvious. The thickness of the underlying 2D wetting layer increases, too. This thickness is deduced from the constant energy of the NP transition of the wetting layer at large coverages. It clearly reflects the reduction of strain with

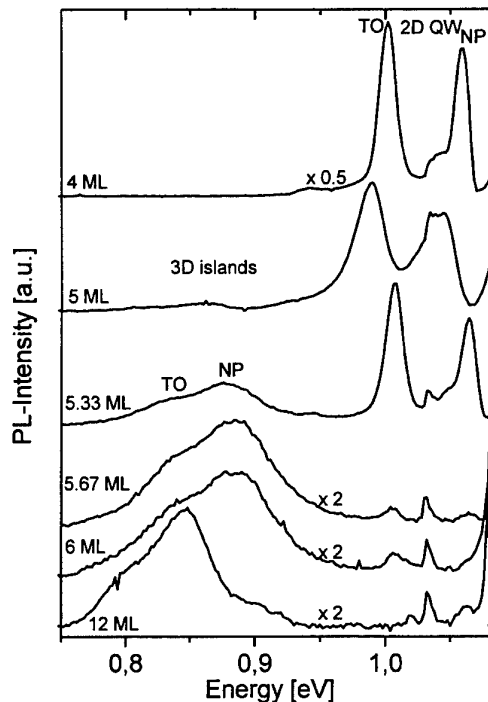


Figure 5. PL spectra of samples with Ge layers of different nominal thickness.

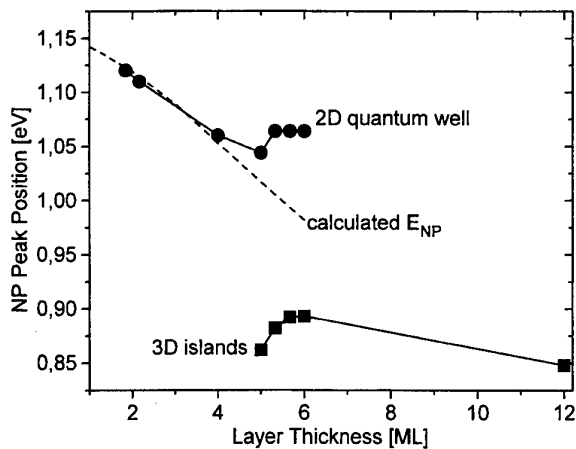


Figure 6. PL energies for the NP transition of 2D quantum well (circles) and islands (squares) in dependence of Ge coverage. The broken curve represents the calculated energies for a 2D quantum well.

decreasing Ge content, and is consistent with the decreasing height of the islands observed with AFM.

A similar behaviour has been found in the InGaAs/GaAs system [20]. The broken calculated curve in figure 6 is based on theoretical results which take confinement, strain and interdiffusion into account [17]. They agree well with the experimental results in the region before island formation occurs. As the PL energies are determined by thickness, composition and strain it is difficult to perform a similar quantitative analysis for the island-related

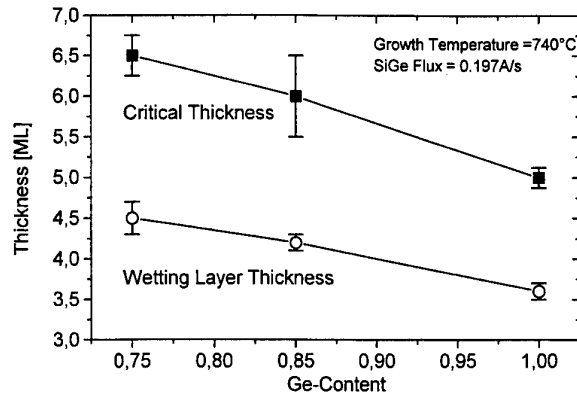


Figure 7. Critical thickness for island formation (squares) and wetting layer thickness (circles) in dependence of the Ge content.

luminescence because the strain distribution in and around the islands is not known. The lateral dimensions of the islands are more than 100 nm and thus much too large to induce lateral confinement. We therefore approximate the islands by a quantum well to get a rough estimate of the Ge content from the luminescence energies. The island heights obtained from TEM images were taken as the thickness of these quantum wells, and the Ge content was varied to fit our calculations [17] to the measured PL energies. If the islands were assumed to be still fully strained, an effective Ge content of only about 40% to 45% in the islands fits the measured values best. This is much lower than expected from a one dimensional diffusion model [21], which would result in about 70% (6 ML) to 90% (12 ML) Ge in the islands for the growth conditions used. For strain relaxed islands, the fit indeed reveals a drastically higher Ge content. So far it has not been possible to solve this problem quantitatively and therefore further work is needed for a microscopic understanding of composition, size and shape of these islands. To achieve pure Ge dots of small sizes, however, it seems unavoidable that the growth temperature must be lowered considerably. This is also evident from figure 8, which shows the PL of dots grown at various temperatures. The PL exhibits a pronounced shift to smaller energies with decreasing growth temperature, indicating an increased Ge content in the islands.

4. Si/SeGe islands/Si pin diodes

First attempts have been made to incorporate Ge-rich islands into electrically active structures. A stack of seven layers was embedded in the intrinsic zone of a pin diode structure as described above. Figure 9 shows the PC spectra of this sample (upper curve) together with a Si pin diode (lower curve) for comparison. Both measurements were performed in normal-incidence geometry on mesa diodes at room temperature. The PC of the pn junction containing Ge islands is about one order of magnitude larger than that of the Si reference sample. Below the energy gap of Si, between 0.95 eV and 1.00 eV, the difference is even two

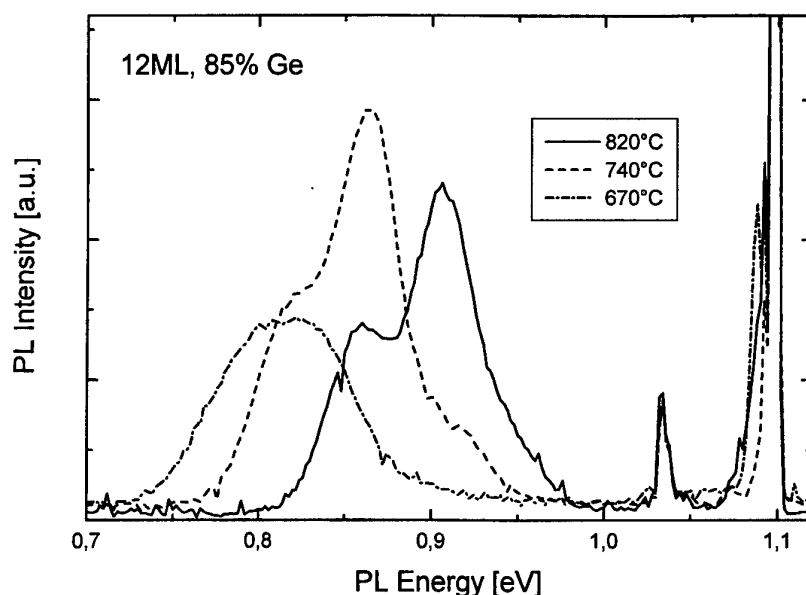


Figure 8. PL spectra of islands formed in nominally 12 ML thick layers of $\text{Si}_{0.15}\text{Ge}_{0.85}$ grown at different temperatures.

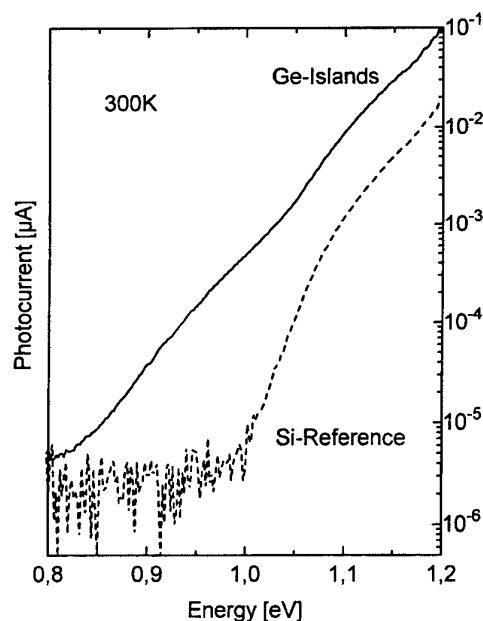


Figure 9. PC spectra of a pin diode with islands embedded in the intrinsic region and of an Si reference diode.

orders of magnitude. The onset of absorption of the Ge-rich islands at 0.825 eV is in excellent agreement with the energy gap obtained from PL and EL data from the same sample. The onset of the Si absorption corresponds very well to the high doping level of the Si layers. The rather good performance of this photodetector structure, together with a low dark current, is regarded as the first evidence that it should be possible to incorporate even multiple layers of Ge islands into device structures without deteriorating the desired optical and electrical properties.

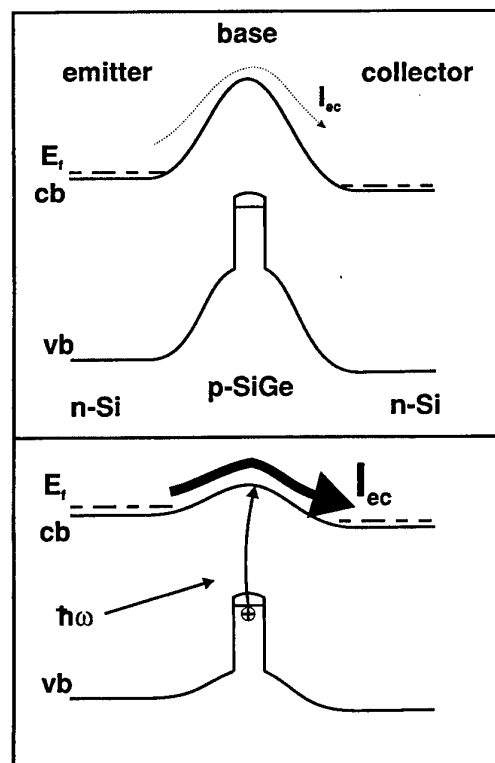


Figure 10. Schematic band diagram of an $n^+pn^+-\text{Si}/\text{SiGe}/\text{Si}$ phototransistor with and without illumination with light. The band offset in the conduction band is neglected.

5. Novel device concepts

Possible applications of Ge-rich islands in Si are based on the reduced energy gap, the band offsets and the band

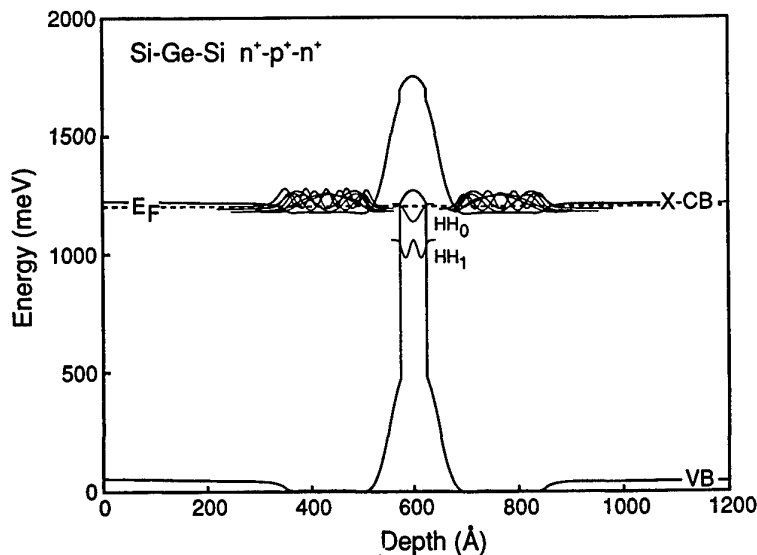


Figure 11. Self-consistent calculation of an $n^+p^+n^+$ -Si/Ge/Si resonant interband tunnelling diode. The parameters have been chosen such that the top of the valence band just reaches the Fermi energy.

alignment, parameters which can be designed by varying the Ge content, the island size and the strain distribution. It was shown more than 10 years ago that the difference in energy gap between Si and SiGe or Ge leads to a large valence band offset and a negligible conduction band offset for pseudomorphic layers on Si substrates [21–24]. The built-in strain lifts the degeneracies of valence and conduction bands and a different strain distribution also sensitively alters the band offsets [24]. For pure Ge on Si an energy gap of only about 0.5 eV is expected for Ge with a valence band offset between 0.7 eV and 0.8 eV. This results in a staggered band line-up with a deep potential well in the valence band. The small energy gap opens the possibility of fabricating infrared-sensitive photodetectors on Si as already demonstrated above. The large valence band offset is used in (nnp) Si/SiGe/Si heterobipolar transistors which have been developed recently for high-frequency applications by various industrial laboratories. Increasing the Ge content in the base region with the ultimate limit of incorporating pure Ge islands in Si leads to further possible applications. As an example, we show in figure 10 a schematic band diagram of a planar doped (nnp) Si/SiGe/Si heterostructure and the relevant emitter-collector current with and without light absorption. The accumulated holes in the Ge-rich base region lower the barrier for the electrons in the conduction band which leads to a strong enhancement and amplification of the emitter-collector current. Trapping of holes, however, may slow down the switch-off behaviour. For small enough Ge islands it might also be possible to apply spectral hole burning for optically writable and perhaps readable information storage by selectively charging individual dots. Electron-hole excitation in a dot leads to trapping of the hole while the electron can move away in an appropriate drift field. Such applications make use of the large valence band offset and the size distribution of the dots, but require

very small dots such that quantization energies are larger than the inhomogeneous broadening of the energy levels.

Apart from optical applications it is also possible to use the large valence band offset in order to enhance interband tunnelling in highly doped pn junctions. The indirect bandgap and the heavy masses of electrons do not make Si an ideal candidate for fabrication of Esaki-type tunnelling diodes. Ge islands embedded at the right position within the pn junction are expected to enhance the tunnelling current considerably by sequential or resonant tunnelling of electrons through empty valence band states in the Ge islands into the p-type region of the Si diode. The tunnelling barrier is reduced drastically and the k -selection rules for the tunnelling processes may also be relaxed due to the rough interface between the Ge islands and the Si. One can even go a step further and try to construct a novel (nnp) Si/Ge/Si resonant tunnelling diode for electrons in the conduction band of Si. In such a symmetric tunnelling structure electrons tunnel from a highly n-doped Si emitter through valence band states of a highly p-doped Ge layer into the conduction band of an n^+ -Si collector. A band diagram of such a structure is shown in figure 11 as calculated self-consistently. In this structure a 5 nm thick Ge layer is assumed to be embedded fully strained in Si. This layer, together with very thin neighbouring Si layers, is p-type doped to a level of $1.6 \times 10^{19} \text{ cm}^{-3}$. The n^+ -Si emitters and collectors are doped to $2 \times 10^{19} \text{ cm}^{-3}$. With such high doping levels one achieves a tunnelling barrier width on each side of the Ge layer of the order of 5 nm and the top of the valence band just reaches the Fermi level. In this calculation only the heavy-hole states in Ge were taken into account. With present epitaxial technologies it should even be possible to fabricate such structures with higher doping levels, although it is not easy to achieve the sharp interfaces and a narrow size distribution of the dots or the homogeneous, but thick enough, Ge layer required for such novel resonant tunnelling devices based on Si.

6. Concluding remarks

Research on self-assembled Ge-rich islands in Si is governed at present by optimization of growth conditions for the achievement of well-defined dots. However, first device structures have already been fabricated and there exist very interesting prospects for future device applications. Some of them, like the resonant tunnelling diodes, are difficult but challenging technological tasks. Such approaches might be a road towards combination of quantum functional devices with conventional Si technology.

Acknowledgments

We gratefully acknowledge fruitful discussions with J F Nützel and M Gail and technical assistance by H Riedl. This work was in part supported by Siemens AG, Munich and by BMBF under contract no 01 M 2953 B 2.

References

- [1] Leonard D, Krishnamurthy M, Reaves C M, DenBaars S P and Petroff P M 1993 *Appl. Phys. Lett.* **63** 3203
 - [2] Moison J M, Houzay F, Barthe F, Leprince L, Andre E and Vatel O 1994 *Appl. Phys. Lett.* **63** 196
 - [3] Grundmann M *et al* 1994 *Phys. Status Solidi b* **118** 249
 - [4] Tanaka T, Komo T, Tsukamoto S, Nishioka M, Uskinow J and Arakawa Y 1995 *Appl. Phys. Lett.* **66** 3663
 - [5] Marzin J Y, Gerard J M, Izrael A, Barrier D and Bastard G 1994 *Phys. Rev. Lett.* **73** 716
 - [6] Nötzel R, Temmyo J, Kamada H, Furuta T and Tamamura T 1994 *Appl. Phys. Lett.* **65** 457
 - [7] Carlsson N, Seifert W, Peterson A, Castrillo P, Pistol M E and Samuelson L 1994 *Appl. Phys. Lett.* **65** 3093
 - [8] Kurtenbach A, Eberl K and Shitara T 1995 *Appl. Phys. Lett.* **66** 361
 - [9] Leon R, Fafard S, Leonard D, Merz J L and Petroff P M 1995 *Appl. Phys. Lett.* **67** 521
 - [10] Hatami F *et al* 1995 *Appl. Phys. Lett.* **67** 656
 - [11] Eaglesham D J and Cerullo M 1990 *Phys. Rev. Lett.* **64** 1943
 - [12] Apetz R, Vescan L, Hartmann A, Diecker C and Lüth H 1995 *Appl. Phys. Lett.* **66** 445
 - [13] Sunamura H, Shiraki Y and Fukatsu S 1995 *Appl. Phys. Lett.* **66** 953
 - [14] Mo Y W, Savage D E, Schwartzentruber B S and Lagally M G 1990 *Phys. Rev. Lett.* **65** 1020
 - [15] Hansson P O, Albrecht M, Dorsch W, Strunk H P and Bauser E 1994 *Phys. Rev. Lett.* **73** 444
 - [16] Schittenhelm P, Gail M, Brunner J, Nützel J F and Abstreiter G 1995 *Appl. Phys. Lett.* **67** 1292
 - [17] Brunner J, Nützel J F, Gail M, Menczgar U and Abstreiter G 1993 *J. Vac. Sci. Technol. B* **11** 1097
 - [18] Oshinowo J, Nishioka M, Ishida S and Arakawa Y 1994 *Appl. Phys. Lett.* **65** 1421
 - [19] Sopanen M, Lipsanen H and Ahopelto J 1995 *Appl. Phys. Lett.* **67** 3768
 - [20] Petroff P M and DenBaars S P 1994 *Superlatt. Microstruc.* **15** 15
 - [21] Van de Walle C G and Martin R M 1986 *Phys. Rev. B* **34** 5621
 - [22] Colombo L, Resta R and Baroni S 1991 *Phys. Rev. B* **44** 5572
 - [23] People R 1986 *IEEE J. Quantum Electron.* **22** 1696
 - [24] Abstreiter G, Brugger H, Wolf T, Jorke H and Herzog H J 1985 *Phys. Rev. Lett.* **54** 2441
- Abstreiter G, Brugger H, Wolf T, Zachai R and Zeller Ch 1986 *Proc. Int. Winterschool in Mauterndorf on 'Two-Dimensional Systems: Physics and New Devices' (Springer Series in Solid State Sciences 67)* ed G Bauer, F Kuchar and H Heinrich (Berlin: Springer) pp 130ff. (See also *Proc. Seventh Int. Winterschool in Mauterndorf on 'Low Dimensional Electronic Systems: New Concepts' (Springer Series in Solid State Sciences 111 (1992).)*)

Optical studies of free-standing single InGaAs/GaAs quantum dots

A Forchel[†], R Steffen[†], T Koch[†], M Michel[†], M Albrecht[†] and
T L Reinecke[‡]

[†] Technische Physik, Universität Würzburg, Am Hubland, D 97074 Würzburg,
Germany

[‡] Naval Research Laboratory, Washington DC, USA

Abstract. Using electron beam lithography and wet chemical etching we have developed single InGaAs/GaAs quantum dots with diameters between 200 nm and 40 nm. The lithographic definition permits the dot shape as well as the dot size to be controlled. In addition to circular dots, which allow the transition between two- and zero-dimensional systems to be studied, we have fabricated rectangular dots which can be used to study the changeover from one- to zero-dimensional structures. Photoluminescence spectra of the dots show a lateral quantization-induced increase of the bandgap energy as well as emission of higher dot states. The experimental results for the lateral confinement energies can be described well by numerical calculations. The average number of electron-hole pairs in the dot can be derived from different features of the spectra. With decreasing dot diameter we observe a continuous reduction of the luminescence linewidths in our structures. This variation is consistent with a decrease of elastic scattering processes by well widths and compositional fluctuations for reduced dot size.

1. Introduction

Optical studies of quantum wires and quantum dots have provided very interesting results regarding the dominant physical effects in quasi one- and zero-dimensional semiconductor nanostructures [1–11]. High-quality structures for optical studies have become available by different technologies. Generally one can distinguish between epitaxy [2–4, 7–10] and lithography [1, 5, 6, 11, 12] based technologies.

Epitaxy-based nanofabrication technologies include, for example, the growth of quantum wire and dot structures on patterned substrates [2, 3]. For the fabrication of quantum dots, self-organization-based technologies have been developed recently. Due to the strong mismatch of the lattice constants of the substrate and the quantum dot material a three-dimensional island growth can be obtained, which results in an *in situ* formation of quantum dots over large areas [7, 8, 10]. The advantages of this technology include the formation of buried dots with high quantum efficiency, high throughput etc. On the other hand, the capability of self-organization techniques to vary shape and size is inherently limited.

Lithography-based technologies provide quantum wires and dots predominantly by the combination of high-resolution electron beam lithography and etching [1, 5, 6, 11, 12]. Here the lateral quantization arises from the confinement of the carriers due to the electron affinity at the etched surfaces. Due to non-radiative recombination

processes at the sidewalls of the structures this method is limited to systems with small surface recombination velocities. In addition process-induced damage, for example by dry etching processes, is known to degrade the quantum efficiency of the nanostructures. Most results on etched wires and dots have been obtained up to now on InGaAs/GaAs and InGaAs/InP structures. As the throughput of electron beam lithography is small, low-dimensional patterns can only be fabricated in small sections of the samples.

On the other hand, lithography-based technologies have several advantages. These technologies permit one to control the size and the separation of the nanostructures in a very flexible way, i.e. by changing the CAD file which controls the electron beam exposure. This is particularly important if the quantum dots or wires are to be integrated in devices. For example, gain coupling in quantum wire or quantum dot arrays is expected to result in low-threshold single-mode lasers, but requires that the nanostructures are fabricated in grids with a periodicity given by the Bragg wavelengths [14, 15]. In addition, the shape of the nanostructures can be controlled by the combined influence of the shape and orientation of the etch mask and the spatial characteristics of the etchant. This permits, for example, the realization of nanostructures with specifically tailored boundary conditions which are of interest for basic physics studies.

Previous studies of lateral quantization effects have been carried out on ensembles of quantum wires or

dots. Due to size fluctuations of a few nanometres the optical spectra are generally strongly inhomogeneously broadened. For a wide variety of epitaxy- and lithography-based fabrication technologies, the increase of the emission linewidths due to inhomogeneous broadening is generally only smaller by a factor of about two than the lateral quantization energies. This broadening often prevents observation of clear lateral sublevel transitions and complicates a direct comparison with experiments. During the last few years single quantum dots have received increasing attention. Because the emission is due to a single nanostructure inhomogeneous broadening effects are avoided. Up to now these studies have been carried out primarily by spatially resolved spectroscopy on samples obtained by self-organization or on thin quantum wells with interface fluctuations [6, 9, 10].

In the present paper we discuss the properties of free-standing InGaAs/GaAs single quantum dots with diameters between 200 nm and 40 nm which have been fabricated by high-resolution electron beam lithography and wet chemical etching. The paper is organized as follows. Section 2 discusses the technology and the experimental set-up used for the characterization. In section 3 experimental results regarding lateral quantization effects and the occupation of the dots by a limited number of electron-hole pairs are discussed. The final section illustrates the potential of the single dots as local probes for the characterization of quantum well structures by focusing on the excitation intensity dependence of the luminescence.

2. Technology and experimental set-up

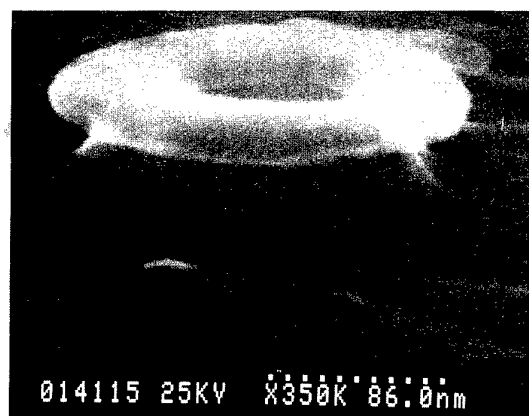
We have developed InGaAs/GaAs single quantum dots by using low-voltage electron beam lithography in PMMA [12, 13]. In this exposure mode, the energy of the electrons is chosen with respect to the resist thickness (here 100 nm) in order to obtain a complete exposure of the resist layer and simultaneously to prevent the electrons from penetrating into the semiconductor substrate. Because the majority of the electrons do not reach the semiconductor material, backscattering from the substrate can be suppressed efficiently. Therefore low-voltage exposures are virtually free of proximity effects. This allows the exposure of strongly varying patterns (for example quantum wire arrays and single dots) with constant exposure doses.

In low-voltage exposures the kinetic energy of the electrons is transferred entirely to the resist. For high-energy exposures, in contrast, most of the electron energy is transferred to the semiconductor substrate, and only a small percentage is used to expose the resist. Thus the sensitivity of the resist at 2 keV for example is larger by a factor of 10 than for exposures at 50 keV [13].

We have used an exposure energy of 2.3 keV to define etch masks for single quantum dots in a 100 nm thick PMMA layer. Al etch masks were obtained by a subsequent lift-off process. The dot masks were etched into the InGaAs/GaAs layers by using $\text{CH}_3\text{OH} : \text{H}_3\text{PO}_4 : \text{H}_2\text{O}_2$ (10:1:1, etch rate 180 nm min⁻¹ at 18°C). After the etching the Al masks were removed in an NaOH solution. Details



(a)



(b)



(c)

Figure 1. SEM micrographs of single InGaAs/GaAs quantum dots prepared by low-voltage electron beam lithography and wet etching. (a) Side view of a circular quantum dot. The dot diameter at the position of the InGaAs layer is 40 nm. (b) Side view of an etched InGaAs/GaAs ring structure. The bright ring structure is the Al etch mask. (c) Top view through the etch mask of a rectangular single dot with a width of 45 nm and a length of 83 nm.

of the fabrication process have been reported elsewhere [12, 13].

Figure 1 illustrates the potential of low-voltage electron beam lithography and wet chemical etching to realize a wide variety of different quantum dot shapes. Figure 1(a) shows a cone-shaped InGaAs/GaAs dot with circular cross section. At the position of the InGaAs layer the dot diameter is about 40 nm. Figure 1(b) shows the side view of a quantum ring structure with an outer diameter of about 150 nm. On top of the semiconductor material the Al etch mask is seen. As shown by figure 1(c), by using a suitable mask shape and orientation a rectangular semiconductor quantum dot may be obtained.

For the optical investigation, the spacing between adjacent single dots was chosen in the range $20\text{ }\mu\text{m}$ to $50\text{ }\mu\text{m}$. This permitted the investigation of individual quantum dots in a standard pumped He cryostat. An Ar^+ laser with excitation power densities in the range 0.1 W cm^{-2} to 10 W cm^{-2} was used as the excitation source. The emission of the single dots was dispersed by a 0.3 m monochromator and detected by a liquid-nitrogen-cooled CCD. The dot spectra were integrated for times ranging from several seconds to about 1 min.

3. Size dependence of optical transitions in free-standing single dots

We have investigated the luminescence of single InGaAs quantum dots as a function of the dot diameter and the excitation power density. Figure 2 shows luminescence spectra of single quantum dots with diameters between 100 nm and 43 nm for an excitation power density of about 1 W cm^{-2} . As will be discussed below, in this excitation power range the dots are occupied by two to three electron-hole pairs at any time. With decreasing dot diameter the luminescence lines shift to higher energies, indicating the influence of the lateral confinement. In addition to the main peak, the spectra of dots with diameters down to 49 nm show a shoulder at high energy which shifts strongly to higher energy as the diameter is decreased.

The change of the transition energies of the lowest (circles) and the first excited state (triangles) of the dots due to lateral quantization is plotted in figure 3 versus the dot diameter. When the dot diameter is reduced from 100 nm to about 40 nm the lateral confinement results in an increase of the dot ground state energy and of the energy of the first excited state by about 6 meV and 9 meV respectively. The full and broken curves represent calculations for the ground and first excited state transition energies respectively. Due to the disc-like shape of the dots the eigenstates are characterized by a radial quantum number n_r and an orbital angular momentum quantum number m in addition to the quantum well subband index n_z . By using Bessel functions the Schrödinger equation in the quantum well plane is solved. The ground state transition is characterized by the quantum numbers $n_r = 0$ and $m = 0$. This implies that, including the spin, this state can be occupied by up to two electron-hole pairs. The first excited state is characterized by $n_r = 0$ and $m = \pm 1$. According to the Pauli principle up to four electron-hole pairs may occupy this level.

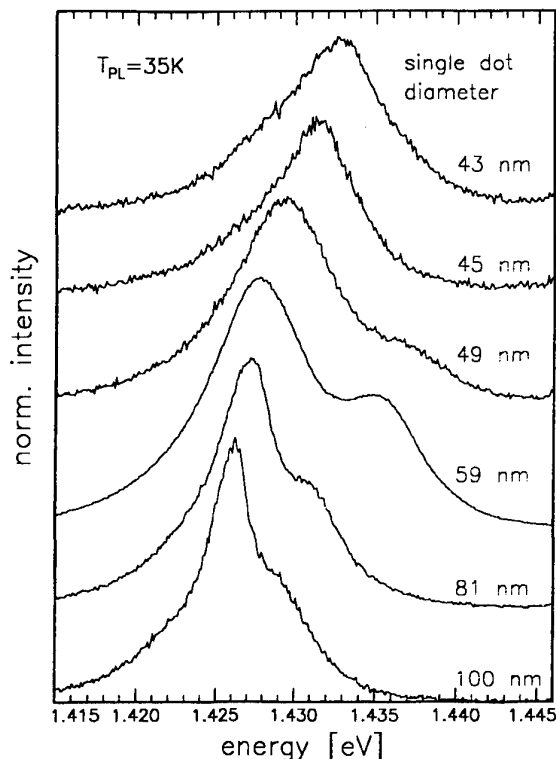


Figure 2. Luminescence spectra of single dots with different diameters for laser excitation intensities which result in the simultaneous occupation of the dots by two to three electron-hole pairs.

The small number of carrier pairs which can be accommodated in the ground state and in the first excited state allows the number of electron-hole pairs which occupy the dot, on the average, to be determined in a simple manner. At very low excitation levels the dot will be unoccupied most of the time. For the remaining time the dot is occupied by a single exciton. As shown by the spectrum at the bottom of figure 4 for the case of a 60 nm dot, the corresponding emission line is quite narrow (FWHM 0.75 meV). If the excitation power is increased (centre trace in figure 4) a pronounced shoulder occurs at low energy of the single exciton luminescence line as well as a weaker feature on the high-energy side. When the excitation power is increased further, the sharp single exciton line vanishes completely and the emission band is dominated by the features at lower and higher energy.

Due to the energetic position below the single exciton transition we assign the low-energy line to the recombination of an electron-hole pair from a completely filled dot groundstate. The two excitons in the groundstate form a biexciton. Due to the attractive interaction between the excitons the emission occurs at lower energy than the single exciton transition. Further electron-hole pairs have to occupy the first excited dot state and therefore give rise to the emission at higher energy.

Figure 5 shows the emission intensity of excitons, biexcitons and of electron-hole pairs in the first excited state versus the exciting laser power. There are pronounced

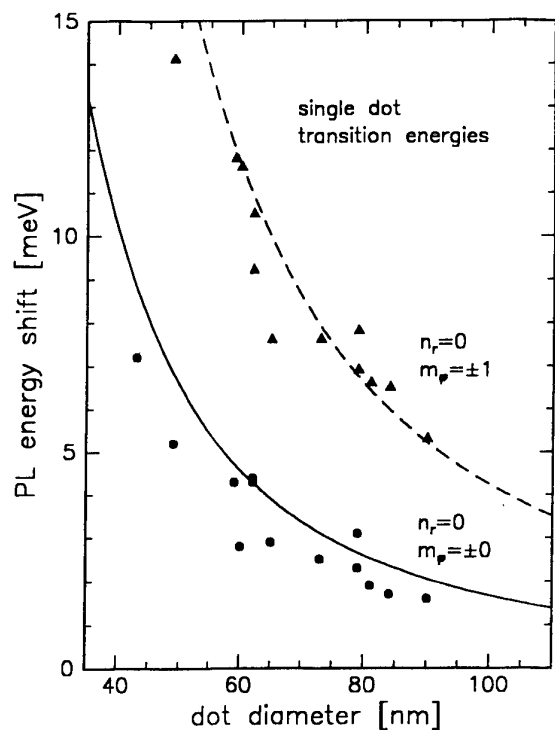


Figure 3. Dot diameter dependence of the energy of the ground state and the first excited state transitions in InGaAs/GaAs single dots: symbols, experimental data; curves, Model calculation.

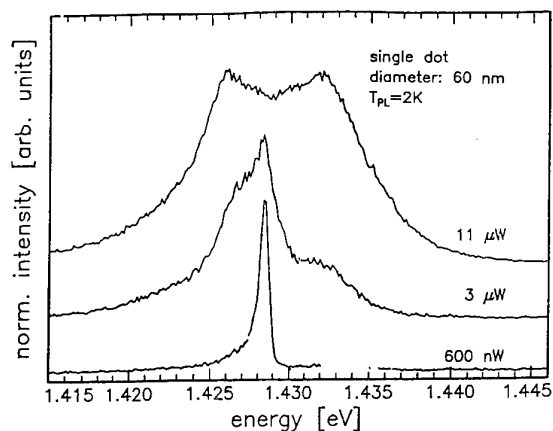


Figure 4. Emission spectra of a 60 nm diameter single quantum dot for different excitation powers [16].

differences between the excitation intensity dependences of the different contributions to the emission. For low excitation power the emission intensity due to recombination of single excitons (bottom part of figure 5) in the dot increases with laser power. Most remarkably, however, a further increase of the excitation power results in a decrease of the emission intensity. For excitation powers on the order of $10 \mu\text{W}$ the emission line due to recombination of single excitons vanishes.

The emission intensity of the biexcitons (centre part

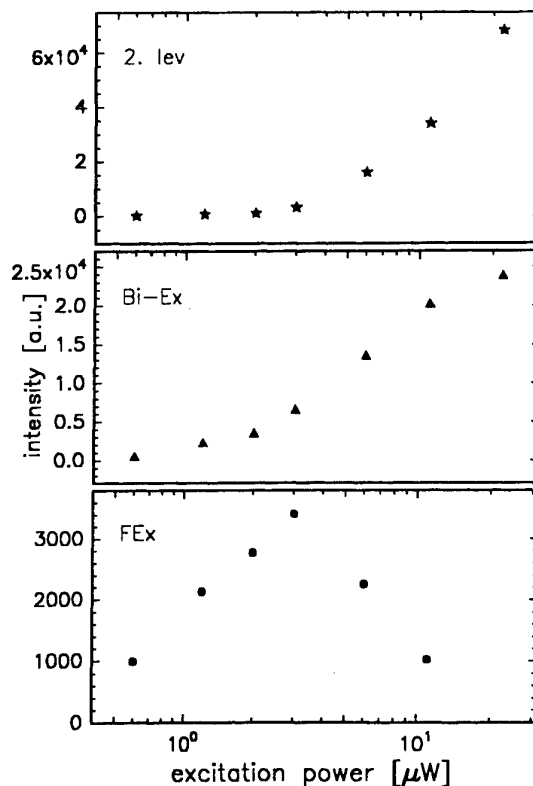


Figure 5. Photoluminescence intensity dependence on the excitation power for the recombination of single excitons (bottom), biexciton emission (centre) and for the emission from the first excited state (top) in an InGaAs/GaAs quantum dot of diameter 60 nm.

of figure 5) shows a nonlinear increase which extends to excitation power levels for which the single exciton line is seen to decrease in intensity. At the highest laser powers investigated the biexciton emission intensity saturates. The saturation implies that the ground state of the dot is occupied at all times by two electron-hole pairs, i.e. a recombining exciton is instantaneously replaced by a new carrier pair generated by the laser excitation. The emission intensity of the single exciton line vanishes in the same excitation range, because the probability of the dot being occupied by a single electron-hole pair goes to zero. We can use the saturation of the emission intensity of the dot ground state to determine the exciton density. For the 60 nm dot investigated here, a completely filled ground state corresponds to a time averaged exciton density of $7 \times 10^{10} \text{ cm}^{-2}$ (top trace in figure 4). Carrier pair densities in dots with partially filled states may be estimated by using this value as a normalization factor for the relative intensities.

The emission intensity of electron-hole pairs in the first excited state (top part of figure 5) increases at higher excitation power levels than the biexciton emission. There is no indication of a saturation in the power range investigated.

We have investigated the halfwidth of the luminescence spectra of free-standing single quantum dots of different

diameters at low excitation (one electron per dot or less at any time) as well as at high excitation (about three electrons per dot at any time) [16]. There are characteristic differences regarding the diameter dependence in both cases. At low excitation intensity the luminescence linewidths decreases continuously when the dot diameter is reduced. At high excitation we observe an increase of the linewidths with decreasing dot diameter.

By using model calculations we relate both trends to elastic scattering processes in the dots. In the case of low excitation, the scattering is most likely associated with well width fluctuations, i.e. to fluctuations in the two-dimensional system. Qualitatively, the magnitude of the fluctuations in a dot decreases as the dot size is reduced, resulting in a decreasing contribution of the fluctuations to the linewidths. The increase of the luminescence linewidths with decreasing dot diameter observed at high excitation, in contrast, is most likely a consequence of the increase of Coulomb interaction in confined geometries. Here a scattering potential is formed by differences in the charge distributions of electrons and holes due to the different penetration of the wavefunctions in the GaAs barrier layers. We have calculated the probability of a scattering process between the degenerate $m = +1$ and the $m = -1$ states of the electron and the hole in a potential formed by two electron-hole pairs in the dot ground state ($m = 0$) as a function of the dot size. The scattering probability, and therefore the respective contribution to the linewidths, increases with decreasing dot size, in qualitative agreement with the experiment. The magnitude of the increase of the linewidths is also consistent with the experimental data [16].

There may be additional contributions to the luminescence linewidths in the dots. For example, by using single dots we have reduced the influence of spatial inhomogeneities in the spectra. Because we integrate over a large number of recombination processes (typically 10^4), however, temporal inhomogeneities such as fluctuations of surface charges in the vicinity of the dots may contribute to the luminescence linewidths, particularly in the case of high excitation.

4. Conclusion

We have investigated the luminescence of free-standing InGaAs/GaAs single quantum dots with diameters between 40 nm and 100 nm. Due to lateral quantization in the etched structures the luminescence of the dots exhibits a clear shift to higher energy with decreasing dot size. In addition to the groundstate transition, we observe for high excitation powers transitions between the first excited states of the dots. The excitation intensity dependence of the emission spectra is used to obtain information on the number of electron-hole pairs occupying the dot. We observe characteristic spectra for the recombination of single excitons, biexcitons, and for the recombination of electron-hole pairs in the first excited state.

In addition to the investigation of lateral quantization effects single quantum dots may be regarded as

spatially resolving probes of the electronic properties of semiconductors. A quantum dot is well defined in all three dimensions. For the present case such a dot corresponds to an InGaAs disc with a thickness of 5 nm and a diameter which is varied down to 40 nm. Because the electron and hole wavefunctions extend through the whole dot, the excitonic recombination may be used to obtain information on compositional fluctuations, carrier interactions etc. Due to the high positional precision, single quantum dots may also be used as local detectors in ambipolar transport experiments. In addition to a large variety of experiments in which the single dots are used as local probes, these structures have particular advantages, because their properties may be modelled exactly by numerical calculations based on the measured geometries of the individual structures. Therefore we are convinced that optical studies of single quantum dots will become powerful tools for the study of a wide variety of electronic properties of semiconductors.

Acknowledgments

We are grateful to F Faller for the supply of high-quality InGaAs/GaAs layers. The financial support of our work by the State of Bavaria and the European Union is gratefully acknowledged. T L Reinecke gratefully acknowledges generous support by the Humboldt Foundation.

References

- [1] Cingolani R, Lage H, Tapfer L, Kalt H, Heitmann D and Ploog K 1991 *Phys. Rev. Lett.* **67** 891
- [2] Naganuma Y, Arakawa Y, Tsukamoto S, Nishioka M, Sasaki S and Miura N 1992 *Phys. Rev. Lett.* **68** 2963
- [3] Kapon E, Hwang D M and Bhat R 1989 *Phys. Rev. Lett.* **63** 430
- [4] Wegscheider W, Pfeiffer L N, Dignam M M, Pinczuk A, West K W, McCall S I and Hall R 1993 *Phys. Rev. Lett.* **71** 4071
- [5] Notomi M, Naganuma M, Nishida T, Tamamura T, Iwamura H, Nojima S and Okamoto M 1991 *Appl. Phys. Lett.* **58** 720
- [6] Clausen E M, Craighead H G, Worlock J M, Harbison J P, Schiavone L M, Florez L and van der Gaag B 1989 *Appl. Phys. Lett.* **55** 1427
- [7] Marzin J Y, Gerard J-M, Izrael A, Barrier D and Bastard G 1995 *Phys. Rev. Lett.* **73** 1138
- [8] Leonard D, Krishnamurthy M, Reaves C M, Denbaars S P and Petroff P M 1993 *Appl. Phys. Lett.* **63** 3202
- [9] Brunner K, Abstreiter G, Böhm G, Tränkle G and Weimann G 1994 *Phys. Rev. Lett.* **73** 1138
- [10] Grundmann M *et al* 1995 *Phys. Rev. Lett.* **74** 4043
- [11] Bayer M, Schmidt A, Forchel A, Faller F, Reinecke T L, Knipp P A, Dremin A A and Kulakovskii V D 1995 *Phys. Rev. Lett.* **74** 3439
- [12] Steffen R, Faller F and Forchel A 1994 *J. Vac. Sci. Technol. B* **12** 3653
- [13] Steffen R, Koch T, Oshinowo J and Forchel A 1996 *Appl. Phys. Lett.* **68** 225
- [14] Lu H, Li G P and Makino T 1993 *Appl. Phys. Lett.* **63** 589
- [15] Luo Y, Nakano Y, Tada K, Hosomatsu H and Iwaoka H 1991 *Appl. Phys. Lett.* **56** 1620
- [16] Steffen R, Forchel A, Reinecke T L, Koch T, Albrecht M, Oshinowo J and Faller F 1996 *Phys. Rev. B* **54** 1510

Towards the optimization of the growth and properties of modulated electron systems on GaAs vicinal surfaces

B Etienne, F Laruelle, Z Wang, L Sfaxi†, F Lelarge, F Petit, T Mélin and A Cavanna

Laboratoire de Microstructures et de Microélectronique (L2M), CNRS, BP 107, 92225 Bagneux Cedex, France

Abstract. The growth of quantum structures with strong in-plane 1D potential modulation (amplitude $V_{pp} \geq E_F$, the Fermi energy of an electron gas in a selectively doped heterojunction) at short scale (periodicity $L \leq \lambda_F$, its Fermi wavelength) relies on the understanding of both the epitaxy and the electronic properties. Some issues such as the morphology of the step array, the organization of the Ga and Al atoms on the surface, and the effect of the electron gas screening or of the tilt of the lateral modulation on the effective potential, are discussed.

1. Introduction

Progress in the fabrication of low-dimensional ($D < 2$) quantum structures using a single technological step such as organized epitaxy requires a good feedback between the study of their electronic properties and the optimization of their growth. This interdisciplinary coupling can be exemplified by the lateral structures made in the nearly lattice matched GaAs/AlAs material system on GaAs vicinal surfaces. There has been much work since Petroff's seminal idea [1] to get to the lateral structures which are discussed here and which exhibit new properties (in optics or transport), through the pioneering demonstrations of fractional lateral superlattices (FLSs) obtained by MOCVD [2] or MBE [3].

2. Surface morphology

The lateral ordering of group III atoms (here Ga and Al) on an array of monomolecular steps (here group III atom terminated steps, so-called A steps, with a ledge oriented along $[1\bar{1}0]$) should be bound to fail completely considering all the detrimental effects summarized in figure 1.

In fact, step pinning by impurities and island nucleation on the terraces can be strongly reduced using an optimized growth procedure and substrate preparation [4]. The terrace distribution has been analysed by AFM (atomic force microscopy) in a dry box after the growth of a GaAs buffer on surfaces misoriented by an angle α ranging from $\sim 0.5^\circ$ to $\sim 2^\circ$ (kinks are not resolved, the lateral resolution

being ~ 4 nm) [5]. In our best growth conditions, this distribution is a Gaussian centred around the nominal value with a standard deviation ΔL (equal respectively to 5.9, 4.1, 2.4 nm for $L = 27.4, 17.0, 8.4$, with $L = a_z / \tan \alpha$ and $a_z = 2.83$ Å). This result could be very well explained in the same way as in Si [6] by a thermodynamical equilibrium model, taking into account $1/L^2$ interactions between steps (originating from entropy and energy terms). But it is clear that in GaAs the kinetics should play a major role in the building up of the periodicity because of the much lower temperature ($\sim 600^\circ\text{C}$ here versus $\sim 900^\circ\text{C}$ in Si) and of the much shorter time scale (0.5–1 ML/s (ML = monolayer) using MBE growth here versus ~ 10 h for Si thermal annealing) used for the surface preparation.

3. Lateral atomic organization

Figure 2(a) depicts an ideal FLS. During the growth, Al adatoms migrating on the terraces can be very efficiently exchanged with Ga atoms underneath [7]. As a result, an AlGaAs alloy is obtained with a lateral modulation of its composition x_{Al} (in the transverse direction y only, if the disorder along the step direction is neglected):

$$x_{\text{Al}}(y) = \bar{x}[1 + \lambda_{\text{at}}(\bar{x}) \cos(2\pi y/L) + \dots]$$

\bar{x} being the mean Al content of the FLS. The parameter λ_{at} , which is called the atomic ordering parameter, depends on \bar{x} but not significantly on L in the investigated range. Because L is small ($L < 35$ nm), very few lateral minibands will be filled for normal charge densities: the electronic properties on the Fermi surface are determined

† Present address: University of Monastir, Tunisia.

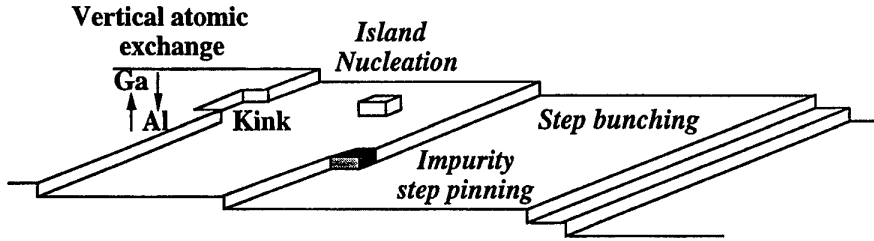


Figure 1. The different phenomena occurring on a vicinal surface, which are detrimental to a perfect lateral atomic ordering. Only those indicated in italics can be avoided (step bunching) or reduced (step pinning, island nucleation).

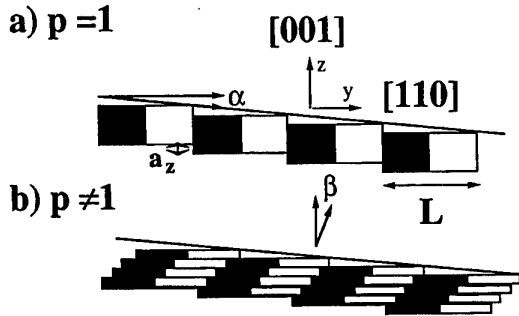


Figure 2. Sketch of an FLS (with perfect lateral atomic ordering) for a coverage ratio $p = 1$ (a) or larger than 1 (b). The drawing is not to scale: the tilt angle β is usually much larger than the misorientation angle α .

by the first Fourier component of the modulation. In the case of a perfect lateral organization, we would get $\lambda_{at}(\bar{x}) = 2 \sin(\pi \bar{x}) / \pi \bar{x} \sim 2$ when $\bar{x} < 1$. In fact only a fraction of the Al atoms are organized, resulting in an alloy lateral composition modulation of amplitude x_{mod} , which gives $\lambda_{at}(\bar{x}) = cx_{mod}/2\bar{x}$ ($c = 1$ for a sine modulation or $4/\pi$ for a square modulation). A summary of the available data on the variations of λ_{at} with \bar{x} is displayed figure 3. As shown, they cannot be explained by a complete lack of Al atom lateral migration, which could be the result of an instantaneous atomic exchange, as proposed in [8]. Another consequence of the Ga/Al exchange is that the lateral atomic organization does not reach its steady-state value $\lambda_{at}(\bar{x})$ right at the beginning of the FLS growth (then it does not even have the step periodicity [9]), but only after some monolayers, shown to be less than 10 [10]. A qualitative model has been proposed to explain these features [4].

The lateral potential modulation is $V_{at}^{e,h}(y) = \Delta E_{e,h} x_{Al}(y)$, where $\Delta E_{e,h}$ is the GaAs/AlAs Γ band offset for electrons or holes. Optical studies [11] have shown that, with a lateral periodicity $L = 32$ nm ($\alpha = 0.5^\circ$) or 16 nm ($\alpha = 1^\circ$), $\lambda_{at}(0.1) \sim 0.20-0.25$ so that the peak-to-peak amplitude, $V_{pp} = 2\lambda_{at}(\bar{x})\bar{x}\Delta E_{e,h}$, can be ~ 50 meV for electrons (half for holes). But because of the finite amplitude of the lateral potential, lower periodicity would not result in a significant modulation of the electronic properties.

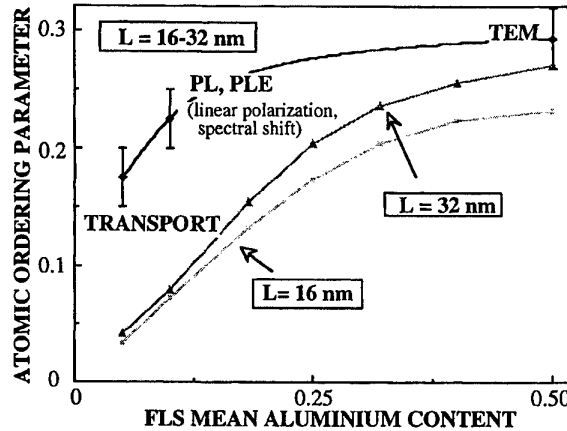


Figure 3. Variation of the lateral ordering parameter λ as a function of the FLS mean aluminium content \bar{x} . Black dots and error bars correspond to the results of the indicated experiments (the dark curve is a guide to the eye) while the lower curves are the result of a calculation (based on the model of [8]) in which no lateral diffusion of the Al atoms is taken into account.

4. 2D calculation of the band structure

Basic solid state physics tells us immediately that the lateral bandgap E_g at $k_y = \pm\pi/L$ and $k_x = 0$ is independent of L and equal to $V_{pp}/2$ (~ 7.3 meV in the conduction band for $(\bar{x}) = 0.05$ and $\lambda_{at} = 0.15$). In a selectively doped heterojunction in which the electrons are transferred into the FLS, this gap is further reduced by screening. Because the periodicity L is comparable to the Fermi wavelength $\lambda_F(\bar{n}_s)$, \bar{n}_s being the mean 2D electron density ($\lambda_F \sim 40$ nm for $\bar{n}_s = 4 \times 10^{11} \text{ cm}^{-2}$), the Thomas-Fermi method would not be very accurate and a 2D calculation of the band structure in the self-consistent Hartree approximation is necessary.

A very efficient method for solving the coupled Schrödinger and Poisson equations has been developed, making extensive use of the fast Fourier transform technique (each wavefunction of the basis is a product of $\zeta_n(z)$ ($0 \leq n < N$), the n th transverse wavefunction of an unmodulated electron gas of density \bar{n}_s in a channel of uniform Al composition (\bar{x}) by a plane wave $\exp(i2\pi q/L)$, with $q = 0, \pm 1, \dots, \pm Q$). The $M \times M$ Hermitian Schrödinger matrix ($M = N(2Q + 1)$) is transformed into a symmetric matrix of size $M \times M$ (not $2M \times 2M$ as easily

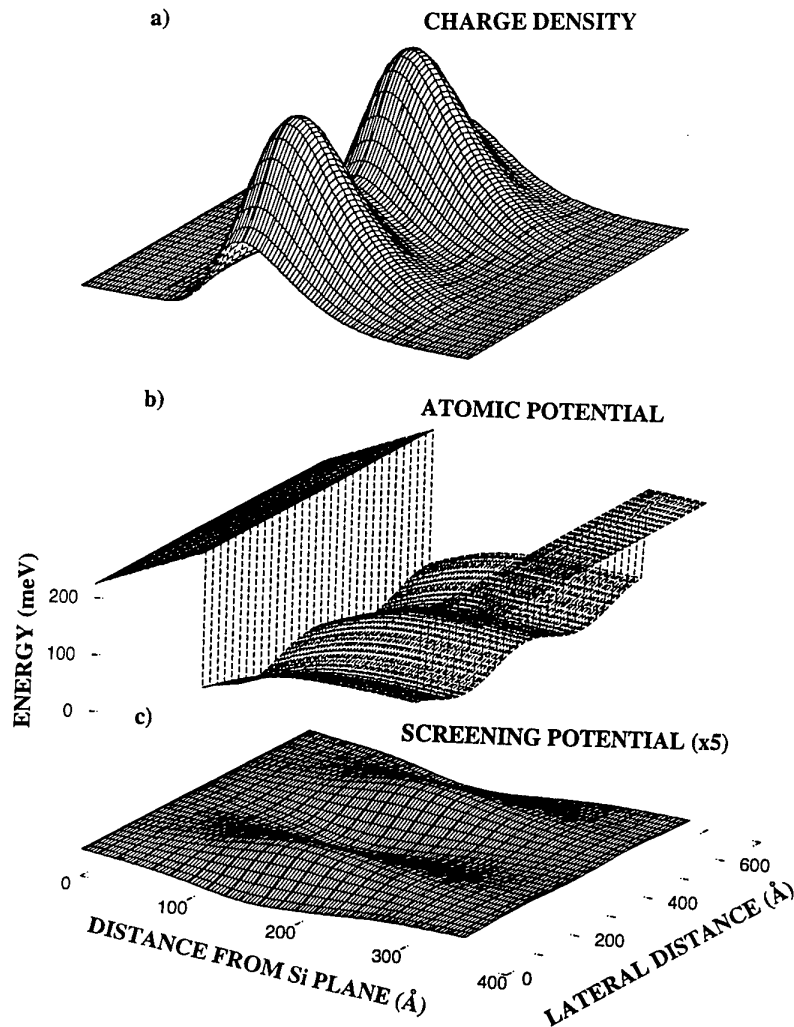


Figure 4. Charge density (a), atomic potential (b) and electrostatic screening potential (c) in a selectively doped heterojunction ($\bar{n}_s = 4.5 \times 10^{11} \text{ cm}^{-2}$) with an FLS ($L = 32 \text{ nm}$ and $\bar{x} = 0.05$). The first Fourier component of the atomic potential is determined experimentally ($\lambda = 0.15$), the charge and the screening potential are then calculated in the self-consistent Hartree approximation.

done!) before being diagonalized completely, in order to save a lot of computing time (it grows as the third power of matrix size). A Green resolvent function $G(y, z, \xi, \eta)$ for the electrostatic potential $V_{el}(y, z)$ obtained from the Poisson equation can be found, assuming as boundary condition that the planar Si doping plane in the spacer is an equipotential plane (ε is the dielectric constant):

$$G(y, z, \xi, \eta) = \frac{1}{\varepsilon} \left\{ \frac{\text{Min}(z, \eta)}{L} + \frac{1}{4\pi} \sum_{m \neq 0} \frac{1}{|m|} \exp\left(i \frac{2\pi m(y - \xi)}{L}\right) \times \left[\exp\left(-\frac{2\pi |m||z - \eta|}{L}\right) + \exp\left(-\frac{2\pi |m||z + \eta|}{L}\right) \right] \right\}$$

(y and ξ are lateral coordinates in the direction orthogonal to the steps, i.e. along [110], while z and η are in the direction of the growth axis, i.e. along [001])

The calculation of $V_{el}(y, z)$ reduces then to a simple integration of the product of this function by the charge density $\rho(\xi, \eta)$. Even if there is some segregation of Si atoms towards the growth front, the effect of the charge modulation is scaled down by a factor $\exp(-4\pi qz/L)$ at a set-back distance z from the heterointerface for the q th Fourier component of the electrostatic potential, so that the boundary condition is well justified ($z \sim 100 \text{ Å}$ or more).

5. Screening and tilt

The screening potential created by the modulated electron density (figure 4) reduces the total potential modulation in the conduction band (CB) but increases it in the valence band (VB). This can be probed by photocreated holes if they are confined in the close vicinity of the degenerate electron gas. We define two parameters λ_{eff}^e and λ_{eff}^h , accounting for

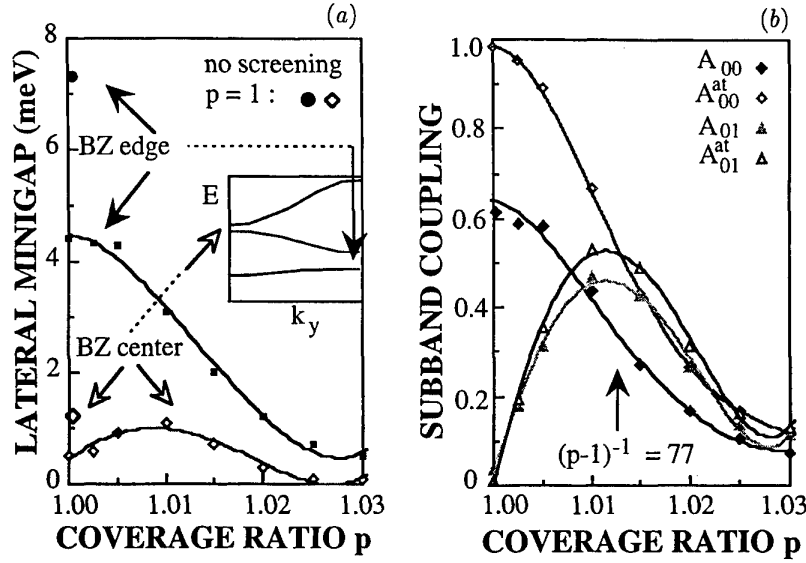


Figure 5. (a) Variation of the first minigap at the Brillouin zone edge and at the zone centre as indicated by the arrows as a function of the coverage ratio p (the magnitude of the gaps for $p = 1$ without screening is also indicated). Insert: sketch of the band structure. (b) Corresponding variations of intrasubband coupling A_{00} and intersubband coupling A_{01} (without screening the coupling coefficients are called A_{ij}^{at}). The latter quantity reaches a maximum for $(p-1)^{-1} = 77$, a value close to the width of the channel (in units of monolayers) [12].

the change of the modulation potential in each band, felt by the charge carriers

$$\lambda_{\text{eff}}^e = (1 - \gamma_e)\lambda_{\text{at}}(\bar{x})$$

$$\lambda_{\text{eff}}^h = (1 + \gamma_h \Delta E_e / \Delta E_h)\lambda_{\text{at}}(\bar{x})$$

$\lambda_{\text{eff}}^e(\lambda_{\text{eff}}^h)$ would give the same lateral Brillouin zone edge minigap E_g in the CB (VB), using a value $\lambda_{\text{at}} = \lambda_{\text{eff}}$ and ignoring screening (which reduces the CB minigap by $1 - \gamma_e$). In GaAlAs, $\Delta E_e / \Delta E_h \sim 2$. Because of the charge spatial separation $\gamma_h < \gamma_e$ (γ_h is nearly equal to γ_e for selectively doped quantum wells of width $W < 80\text{--}100 \text{ \AA}$). With $L = 32 \text{ nm}$, $W = 20 \text{ nm}$ and $\bar{n}_s = 4.5 \times 10^{11} \text{ cm}^{-2}$, we get $\gamma_e = 0.32$ and $\gamma_h = 0.17$. Therefore for $\lambda_{\text{at}} = 0.15$, $\lambda_{\text{eff}}^e = 0.10$ and $\lambda_{\text{eff}}^h = 0.20$ are expected. This explains why different experiments may lead to values for λ which seem in complete disagreement if screening is not taken into account: a measurement of the linear polarization ratio of the radiative recombination between degenerate electrons and photocreated holes is determined by the properties of the VB states [11] and therefore gives λ_{eff}^h , while electrical measurements reflect the CB properties and give λ_{eff}^e .

In addition, such an FLS can be easily tilted (figure 2(b)) because of some inaccuracy in the group III atom flux calibration (it is difficult to get better than 1% in MBE) or of the lateral gradient of these fluxes (the temporal variations are usually negligible during the duration of the FLS growth). The coverage ratio p is defined as the exact amount of matter (in units of monomolecular monolayers, i.e. of a thickness equal to a_z) deposited by a nominal monolayer. It is independent of L , whereas this is not the case for the tilt angle β , which is therefore not a very meaningful parameter. If $p \neq 1$, the lateral potential

modulation depends also on the coordinate z along the growth axis: we now have $V_{\text{at}}^{e,h}(y, z)$. This reduces again the bandgap as shown for electrons in figure 5(a). For states originating from the first electrical subband ($n = 0$), the corresponding effective atomic potential would be $(\varphi(p))$ is a phase term) [12]

$$V_0^{e,h}(y) = \langle \zeta_0^{e,h}(z) | V_{\text{at}}^{e,h}(y, z) | \zeta_0^{e,h}(z) \rangle \\ = \bar{x} \Delta E_{e,h} [1 + \lambda_{\text{at}}(\bar{x}) A_{00}^{e,h} \cos(2\pi y/L + \varphi(p))].$$

The Brillouin zone edge minigap in the CB follows exactly this intrasubband coupling term A_{00}^e (figures 5(a) and (b)). We can define an effective value for λ , called $\lambda_1(p)$, which would give the same bandgap E_g with $p = 1$ as the tilted FLS having $p \neq 1$. The ratio $\lambda_1(p)/\lambda_{\text{at}}(p)$ would again follow the curves of figure 5. It can be concluded that it is required to get $p = 1 \pm 0.01$ in order to avoid a significant reduction in the effective lateral potential modulation. But the tilting effect also changes the symmetry of the lateral potential, inducing intersubband coupling $A_{nn'}(p)$ with $n \neq n'$ [12]. These non-diagonal terms are equal to zero for $p = 1$. This would be in fact exactly true only if the atomic modulation were to go on into the barrier (this not the case in our structures but the electron density is negligible there). It has to be noted also that the electrostatic screening potential, being dependent on the z coordinate, also induces some intersubband coupling (using the total potential in the relation written above) even if the atomic potential is not tilted itself (figure 5(b)). The $A_{nn'}(p)$ go through a maximum for a value of $p \neq 1$. This effect (for $n = 0$ and $n' = 1$) can be revealed by the dependence on p of the Fermi edge singularity observed in low-temperature optical studies [13]. Another effect of the interband coupling is that the bandgap at the centre of the Brillouin zone is much less

reduced than the bandgap E_g at the Brillouin zone edge, when the second electric subband is not far away (this is the case when $L \sim 32$ nm) (figure 5(a)).

6. Conclusion

With a progressive understanding of all these effects, we have been able to get reproducible and well understood data in quantum structures having such an FLS in the active region. In a first step, we have found expected properties: spectral shift (with respect to an unmodulated structure) and linear polarization of the photoluminescence and PL excitation spectroscopy [11], anisotropy of the electrical conductivity, magnetoresistance for current flow in the direction perpendicular to the steps and magnetic breakdown [14]. More recently, in a second step, we have been able to observe new properties due to the strong, periodic and low-disordered lateral modulation: a non-monotonic variation of the Fermi edge singularity with the tilt [13] and a transition in the spin polarization of the Landau bands at high magnetic field [15]. All these phenomena can be analysed with fairly consistent values for $\lambda_{at}(\bar{x})$, provided that screening (calculated here in the Hartree approximation) and tilt, if any, are properly taken into account.

Much work has still to be done. The building up of the step periodicity and of the lateral ordering of the Ga and Al atoms are not yet really understood. This would be helpful in order to know more about the ultimate possibility of this type of organized growth and to assess more quantitatively the importance of the fluctuations.

Acknowledgments

We acknowledge partial support by DRET and by the EEC (HCM network GOODS).

References

- [1] Petroff P M, Gossard A C and Wiegman W 1984 *Appl. Phys. Lett.* **45** 620
- [2] Fukui T and Saitoh H 1988 *J. Vac. Sci. Technol. B* **6** 1373
- [3] Gaines J M, Petroff P M, Kroemer H, Simes R J, Geels R S and English J H 1988 *J. Vac. Sci. Technol. B* **6** 1378
- [4] Etienne B, Laruelle F, Bloch J, Sfazi L and Lelarge F 1995 *J. Crystal Growth* **150** 336
- [5] Lelarge F, Lebreton C and Wang Z Z 1995 Presented at STM 95, Aspen (unpublished)
- [6] Bartelt N C, Einstein T L and Williams E D 1990 *Surf. Sci. Lett.* **240** L591
Alfonso C, Bermond J M, Heyraud J C and Métois J J 1992 *Surf. Sci.* **262** 371
see also Villain J and Pimpinelli A 1995 *Physique de la Croissance Cristalline* edn Eyrolles and CEA (in French)
- [7] Moison J M, Guille C, Houzay F, Barthe F and van Rompay M 1989 *Phys. Rev. B* **40** 6149
- [8] Lorke A 1995 *Low Dimensional Structures Prepared by Epitaxial Growth or Regrowth on Patterned Substrates* ed K Eberl, P M Petroff and P Demeester (Dordrecht: Kluwer) p 139
- [9] Bloch J, Bockelman U and Laruelle F 1994 *Solid-State Electron.* **37** 529
- [10] Lelarge F, Laruelle F and Etienne B Euro-MBE 95, Granada, Spain (unpublished)
- [11] Bloch J, Bockelman U and Laruelle F 1994 *Europhys. Lett.* **28** 501
- [12] Lelarge F, Mélin T and Laruelle F 1995 *Semiconductor Heteroepitaxy (Growth, Characterization and Device Applications)* ed B Gil and R L Aulombard (Singapore: World Scientific) p 339
- [13] Mélin T and Laruelle F submitted
- [14] Sfazi L, Petit F, Lelarge F, Cavanna A and Etienne B 1995 Presented at MSS-7, Madrid (Proceedings to appear in *Solid State Electron.*); 1995 Presented at EP2DS-11, Nottingham (Proceedings to appear in *Surf. Sci.*)
Etienne B 1995 *Ann. Physique* suppl **20** C2-235 (in French)
- [15] Petit F, Sfazi L, Lelarge F, Cavanna A and Etienne B submitted

Correlated bilayer electron states

M Shayegan, H C Manoharan, Y W Suen, T S Lay and
M B Santos

Department of Electrical Engineering, Princeton University, Princeton, NJ 08544,
USA

Abstract. We report magnetotransport measurements in a wide quantum well as the electron charge distribution is tuned from a single-layer through an interacting bilayer configuration to weakly coupled parallel layers. The system exhibits a remarkably rich set of correlated *bilayer* states including unique fractional quantum Hall states at even-denominator fillings and insulating phases which are consistent with pinned, bilayer Wigner crystal states.

1. Introduction

The introduction of an additional degree of freedom can have a profound effect on the many-body ground states of the two-dimensional electron system (2DES) at high magnetic fields B . For example, the addition of a spin degree of freedom stabilizes the particular spin-unpolarized fractional quantum Hall effect (FQHE) observed at lower B [1, 2], while substantially increasing the layer thickness (thus introducing an additional spatial degree of freedom) leads to a weakening and eventual collapse of the FQHE [3, 4].

Here we report magnetotransport measurements on an electron system (ES) confined in a wide GaAs quantum well, which can be tuned from a single-layer-like (albeit thick) system to a bilayer system by increasing the electron density N in the well [5]. This evolution with N and, in particular, the transition to a bilayer system where *interlayer* as well as *intralayer* interactions are dominant, has a dramatic effect on the correlated states of the ES, as manifested in the magnetotransport data [6–10].

Figures 1 to 3 provide examples for an ES in a 750 Å wide GaAs quantum well. In certain ranges of N , there are well developed FQH states at the *even-denominator* fillings $\nu = 1/2$ [6–8]† and $3/2$ [8] which have no counterparts in standard 2DESs in single heterostructures. (Note that $\nu = 1/2$ is the total filling for the system; it corresponds to $1/4$ filling for each layer). Figure 2 shows that at $N = 1.26 \times 10^{11} \text{ cm}^{-2}$, insulating phases (IPs), which are *re-entrant* around the $\nu = 1/2$ FQH state, develop. The data in figure 2 have a remarkable resemblance to the IPs observed in very high-quality, standard, GaAs 2DESs [12] except that here the IP is re-entrant around the much higher filling $\nu = 1/2$ rather than $\nu = 1/5$! Figure 3 demonstrates yet another surprising aspect of the QHE in this system. Here the Arrhenius plots of resistance at $\nu = 1$ are very unusual: the activated behaviour of R_{xx} versus $1/T$ starts rather abruptly below a temperature

† The $\nu = 1/2$ FQH state is also observed in bilayer electron systems in *double* quantum wells [11].

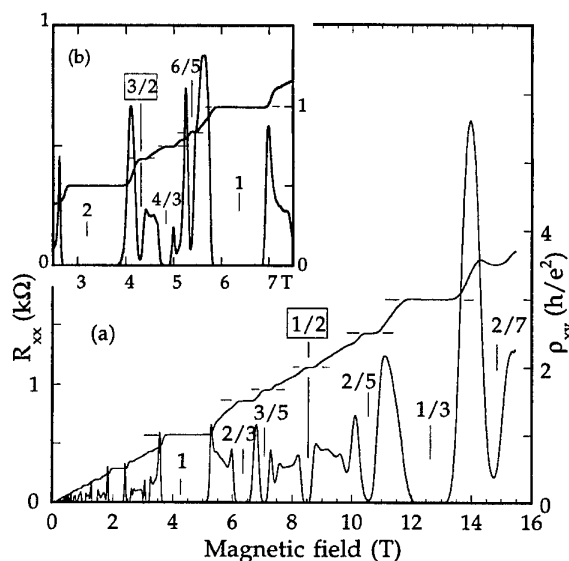


Figure 1. Magnetotransport data, taken at $T \approx 30 \text{ mK}$, for a 750 Å wide well with (a) $N = 1.03 \times 10^{11} \text{ cm}^{-2}$ and (b) $N = 1.55 \times 10^{11} \text{ cm}^{-2}$, showing well developed *even-denominator* FQH states at $\nu = 1/2$ and $3/2$.

T^* which is much smaller than the deduced QHE gaps ($\approx 20 \text{ K}$) and, even more surprisingly, is strongly N -dependent.

The evolution of the magnetotransport data in this system as a function of N , and the intriguing interplay between the incompressible liquid states (including the $\nu = 1/2$ FQHE) and the IP, which displays behaviour profoundly different from any observed in a standard 2DES, are the subject of this paper. We first discuss the FQHE in this system and then focus on the IP, which we associate with a pinned, *bilayer* Wigner crystal with intralayer and interlayer correlations. Finally, we summarize our results for the $\nu = 1$ QHE in this system [9] and suggest that its very unusual dependence on T and N may be indicative of an additional *finite-temperature* transition, from a QHE

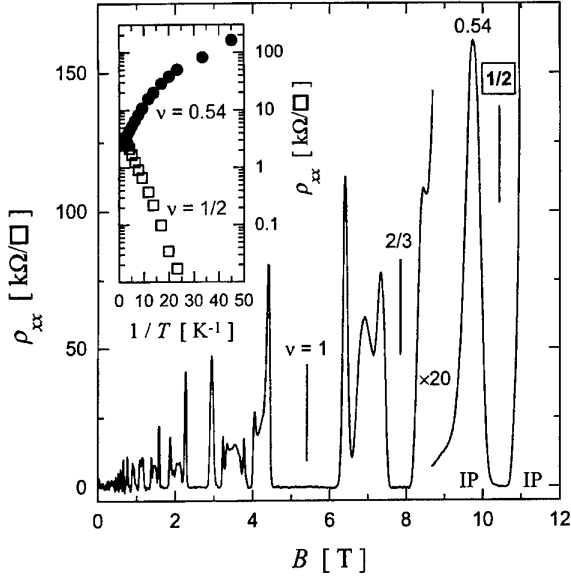


Figure 2. Data for the same sample as figure 1 but with $N = 1.26 \times 10^{11} \text{ cm}^{-2}$. Here we observe an insulating phase re-entrant around the $\nu = 1/2$ FQH state. The inset shows the temperature dependence of resistivity: at $\nu = 1/2$, ρ_{xx} vanishes as $T \rightarrow 0$ indicative of a FQH state while at slightly higher and lower ν , ρ_{xx} shows an insulating behaviour as it diverges with decreasing T .

state to a compressible state, which is unique to interacting bilayer systems.

2. Electron system in a wide, single quantum well

The ES in a wide GaAs quantum well of width $\sim 1000 \text{ \AA}$ [5] is a particularly interesting one. At low N the electrons occupy the lowest electric subband and have a single-layer-like (but rather ‘thick’ in the z -direction) charge distribution (figure 4). As more electrons are added to the well, their electrostatic repulsion forces them to pile up near the well’s sides and the resulting electron charge distribution appears increasingly bilayer-like. A relevant parameter that quantifies this evolution is the energy difference between the symmetric and antisymmetric subbands (Δ_{SAS}) which is a measure of the coupling between the two layers. Also relevant is the interlayer distance, defined by the parameter d as shown in figure 4. A crucial property of the ES in a wide quantum well is that, for a given well width, both Δ_{SAS} and d depend on N : increasing N makes d larger and Δ_{SAS} smaller so that the system can essentially be tuned from a (thick) single-layer-like ES at low N to a bilayer ES by increasing N (figure 4). This evolution with density plays a decisive role in the properties of the correlated electron states in this system.

Experimentally, we control both N and the charge distribution symmetry in the ES via front- and back-side gates, and by measuring the occupied subband electron densities from Fourier transforms of the low- B magnetoresistance oscillations. The Δ_{SAS} data of figure 4 were obtained from such measurements on ‘balanced’ states, i.e. the gates were tuned to preserve symmetric

charge distributions in the well. The remarkable agreement of the data with the self-consistent calculations, and the controlled variation of Δ_{SAS} and d with N , attest to the excellent tunability of the ES in a wide quantum well.

Besides this tunability, the bilayer ES in a wide GaAs well has another great advantage over its counterpart in a double-quantum-well (DQW). Here the effective ‘barrier’ separating the two electron layers is GaAs while in a DQW the barrier is AlGaAs or AlAs. The purity of GaAs grown in a molecular beam epitaxy chamber is typically higher than that of AlGaAs or AlAs (because of the high reactivity of Al and its sensitivity to impurities). Moreover, the AlGaAs or AlAs barrier in a DQW introduces additional interfaces which often adversely affect the quality of the ES, for example because of the additional interface roughness scattering.

3. Evolution of the FQHE in a wide quantum well

Light is shed on the origin of the FQH states at $\nu = 1/2$ and other fillings in a wide quantum well by examining the competition between (i) Δ_{SAS} , (ii) the in-plane correlation energy $Ce^2/\epsilon l_B$ (where C is a constant ~ 0.1 and $l_B \equiv (\hbar/eB)^{1/2}$ is the magnetic length), and (iii) the interlayer Coulomb interaction $\sim e^2/\epsilon d$. To quantify behaviour it is useful to construct the ratios $\gamma \equiv (e^2/\epsilon l_B)/\Delta_{\text{SAS}}$ and $(e^2/\epsilon l_B)/(e^2/\epsilon d) = dl_B$. As N is increased, γ increases since both Δ_{SAS} and l_B (for an FQH state at a given ν) decrease, and d/l_B increases. When γ is small, the system should exhibit only ‘one-component’ (1C) FQH states (standard single-layer odd-denominator states) constructed from only the symmetric subband, while for large γ the in-plane Coulomb energy becomes sufficiently strong to allow the antisymmetric subband to mix into the correlated ground state to lower its energy and a ‘two-component’ (2C) state ensues. These 2C states, constructed out of the now nearly degenerate symmetric and antisymmetric basis states, have a generalized Laughlin wavefunction of the form [13–15]

$$\Psi_{mnn}^{\nu} = \prod_{i,j} (u_i - u_j)^m \prod_{i',j'} (w_{i'} - w_{j'})^m \prod_{i'',j''} (u_{i''} - w_{j''})^n \times \exp \left(- \sum_i \frac{|u_i|^2}{4l_B^2} - \sum_{i'} \frac{|w_{i'}|^2}{4l_B^2} \right)$$

where in a more intuitive pseudo-spin or two-layer language, u_i and w_i denote the pseudo-spin or the complex 2D coordinates of an electron in the two layers. The integer exponents m and n determine the intralayer and interlayer correlations, respectively, and the total filling factor for the Ψ_{mnn}^{ν} state is $\nu = 2/(m+n)$.

Now the 2C states described by Ψ_{mnn}^{ν} come in two classes. For large d/l_B , the system behaves as two independent layers in parallel, each with density $N/2$. FQH states in this regime therefore have even numerator and odd denominator. An example is the $\Psi_{330}^{2/3}$ state which has a total filling of $2/3$ ($1/3$ filling in each layer). For small enough d/l_B , on the other hand, the interlayer interaction can become comparable to the in-plane interaction and a fundamentally new kind of FQH state becomes possible. Such a state has strong interlayer correlation and can be at even-denominator ν ; a special example is the $\Psi_{331}^{1/2}$ state

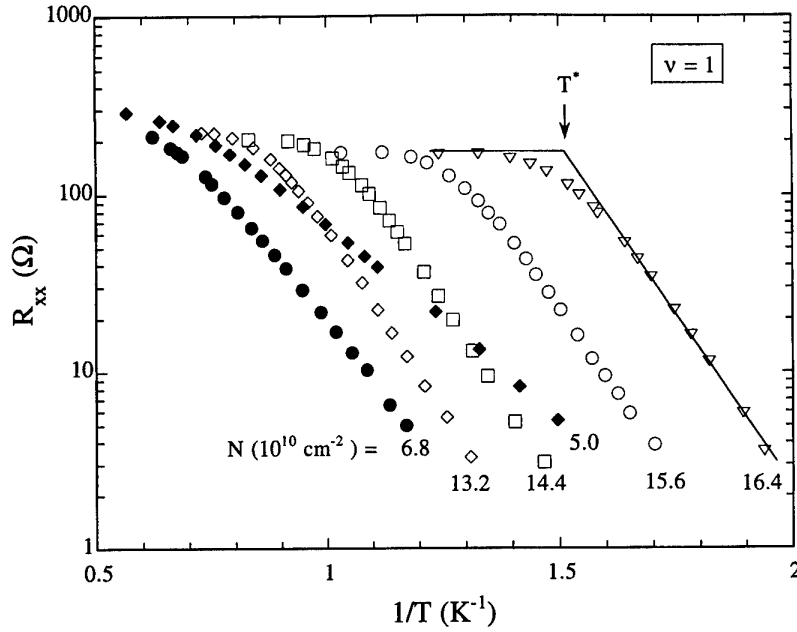


Figure 3. Arrhenius plots of R_{xx} at $\nu = 1$ versus $1/T$ for the sample of figures 1 and 2 for different N . For $N > 10 \times 10^{10} \text{ cm}^{-2}$, the activated behaviour of R_{xx} ends abruptly above a temperature T^* which strongly depends on N .

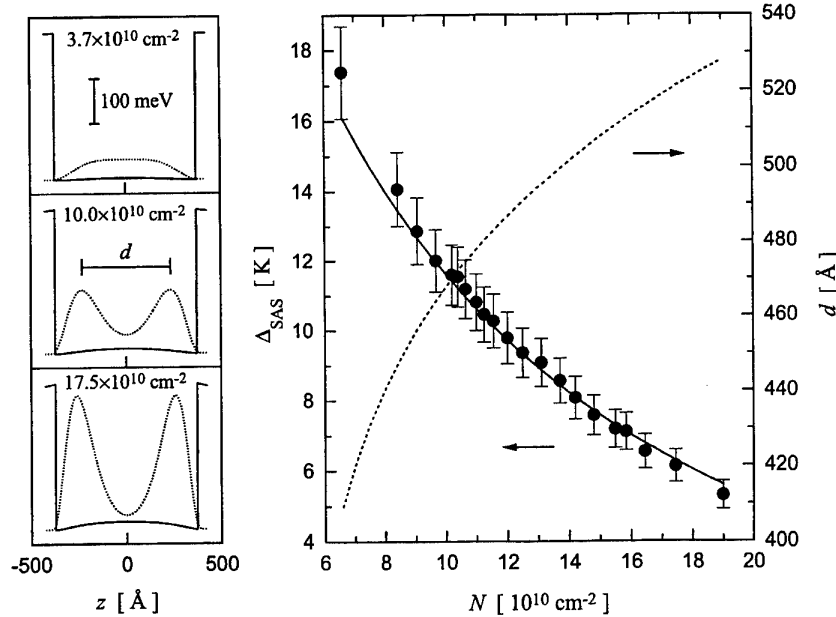


Figure 4. The evolution of an electron system in a 750 \AA wide single quantum well as a function of total electron density N . On the left the results of Hartree-Fock simulations for the self-consistent conduction band potentials (full curves) and charge distributions (dotted curves) are shown. On the right the calculated Δ_{SAS} (full curve) is compared with the measured Δ_{SAS} (full circles); also shown is the calculated layer separation d .

with $\nu = 1/2$. (The physics here is also related to that of the $\nu = 5/2$ FQH state which is observed in standard 2DESs at relatively low B where the Zeeman energy is small, thus providing a spin degree of freedom [1].)

A careful study of the evolution of the FQH states in a wide quantum well as a function of N and the symmetry of the charge distribution reveals that this evolution is compatible with the above picture [8]. In particular,

the quasiparticle excitation gaps of several FQH states determined via thermal activation measurements depend on γ as shown in figure 5. As expected, increasing γ suppresses 1C states (such as $3/5$) and enhances 2C states (such as $4/5$). Two states, $\nu = 2/3$ and $\nu = 4/3$, undergo a 1C to 2C phase transition as γ is increased. The critical point for this transition, $\gamma \approx 13.5$, is consistent with the ratio of the in-plane correlation energy and Δ_{SAS}

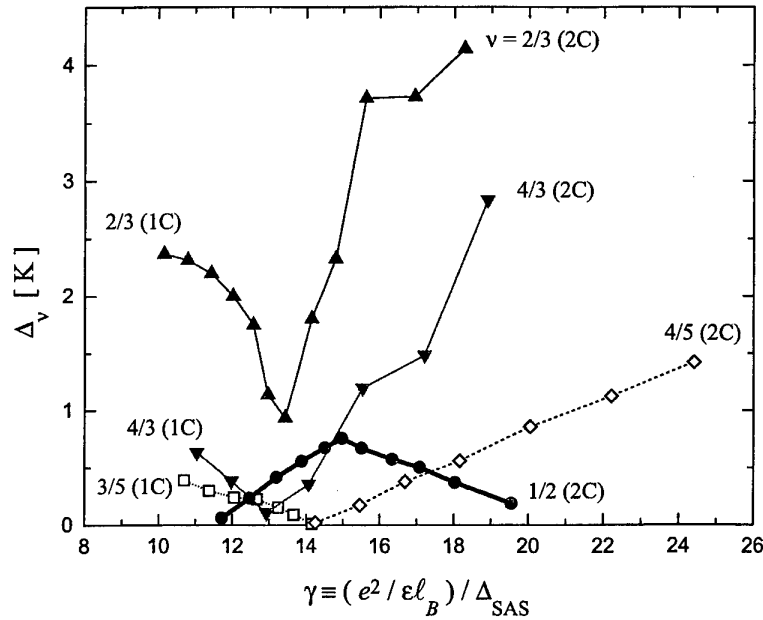


Figure 5. Measured energy gaps Δ_v of several FQH states versus γ . The number of components (C) in each state is marked in parentheses. As γ increases, the 1C to 2C transition is observed at $\gamma \simeq 13.5$. The $\nu = 1/2$ state is correlated both in the plane and between layers, and thus exists only within a finite range of γ .

($\sim 0.1(e^2/\epsilon l_B)/\Delta_{SAS} = 0.1\gamma$) being of the order of unity, and matches the point where the energy gaps of other 1C and 2C states emerge from zero. Surrounding this point is a region where the $\nu = 1/2$ FQH liquid stabilizes. Note that since this is a 2C state which possesses interlayer correlation in addition to intralayer correlation, it exists only within a finite range of γ (the 2C $\nu = 2/3$ and $4/3$ states are simply $1/3$ and $2/3$ states in parallel layers).

4. Evolution of insulating phases

Concurrent with the evolution of the FQH states in this wide quantum well, we observe an IP which moves to *higher* ν as N is increased. Our data are summarized in figure 6 where ρ_{xx} at base T is plotted versus $\nu^{-1} \propto B$ for several representative N . Experimentally, the IP is identified by a resistivity that is both large ($\rho_{xx} > h/e^2 \approx 26 \text{ k}\Omega \square^{-1}$, the quantum unit of resistance) and strongly T dependent. For very low N , the IP appears near $\nu = 1/5$, while at the highest N there is an IP for $\nu \lesssim 1/2$. The IP observed in the intermediate density range ($10.4 \times 10^{10} < N < 14.2 \times 10^{10} \text{ cm}^{-2}$) is most remarkable as it very quickly moves to larger ν with small increases in N (see, for example, traces B, C and D in figure 6); along the way, it also shows *re-entrant* behaviour around well developed FQH states at $\nu = 2/7$ (trace B), $\nu = 1/3$ (traces C and D) and $\nu = 1/2$ (trace E). Then, as N increases past this point, the IP begins to move in the opposite direction to lower ν (trace F). The data shown in figures 2 and 5 for $N = 12.6 \times 10^{10} \text{ cm}^{-2}$ bear a striking resemblance to the IP observed re-entrant around $\nu = 1/5$ in low-disorder, single-layer 2DESs, generally interpreted as a pinned Wigner solid (WS) [12]; here, however, the IP is re-entrant around

the bilayer $\nu = 1/2$ FQH state, with the re-entrant peak reaching the prominently high filling of $\nu = 0.54$.

The IPs presented in figure 6 cannot be explained by single-particle localization. First, in the case of standard, single-layer 2DESs it is well known that as N is lowered, the quality of the 2DES deteriorates and the sample shows a disorder-induced IP at progressively larger ν [16]. This is opposite to the behaviour observed here: as N decreases from 10.9×10^{10} to $3.7 \times 10^{10} \text{ cm}^{-2}$, the quality worsens as expected (for example mobility decreases monotonically from 1.4×10^6 to $5.3 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$) but the IP moves to *smaller* ν . Second, the observation of IPs which are re-entrant around *correlated* FQH states, and particularly around the very fragile $\nu = 1/2$ state [8], strongly suggests that electron interactions are also important in stabilizing the IP.

To illustrate that the behaviour of this IP is indeed consistent with the WS picture, it is instructive to first examine the three main re-entrant peaks in figure 6 (from traces B, D and E), which appear at $\nu = 0.30, 0.39$ and 0.54 for the IPs surrounding the $\nu = 2/7, 1/3$ and $1/2$ FQH states respectively. The values of γ at these peaks are respectively $16.9, 16.3$ and 16.5 . The peak positions span a large region of ν , and yet the associated γ are remarkably similar. Moreover, at this value of $\gamma \approx 16.5$, interlayer interactions are clearly important as this point is straddled by the 2C $1/2$ state in figure 5.

The construction of a phase diagram for the observed IPs facilitates a clear correlation between the IP evolution, the 1C to 2C transition, and the development of the $\nu = 1/2$ liquid. We have examined the T -dependence of ρ_{xx} at many N over the entire measurement range described earlier. We observe T -independent critical points, where $d\rho_{xx}/dT$ changes sign, in the high- B ($\nu < 2/3$) regime

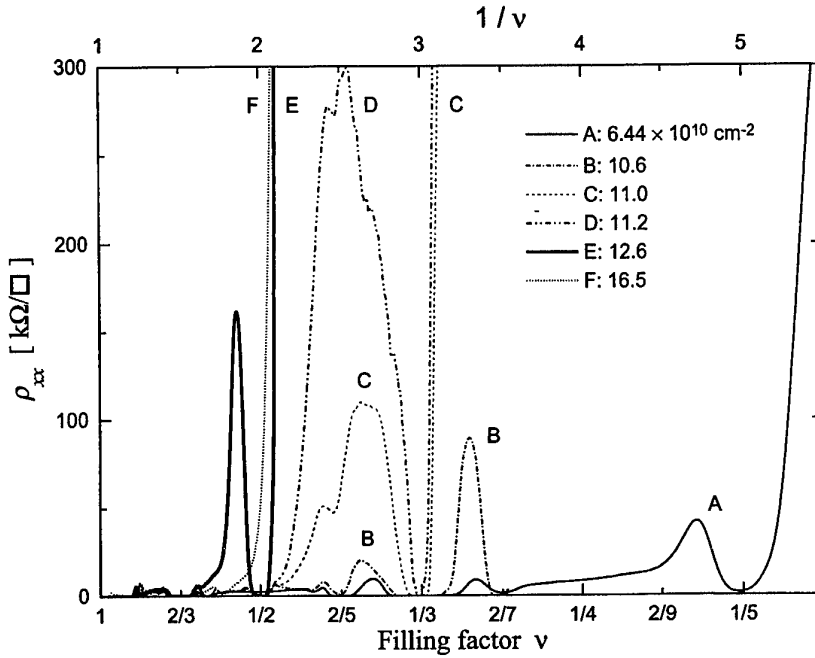


Figure 6. Evolution of the IP at $T \approx 25$ mK. As N is increased, the IP moves quickly in to higher ν , becoming re-entrant around several FQH states along the way, until it stabilizes around the $\nu = 1/2$ bilayer state (bold trace E). As N is further increased from this point, the IP reverses direction and begins to move out toward lower ν .

when ρ_{xx} becomes large. The B fields of these critical points may be used to demarcate the boundary between the IP and non-insulating states. As there is now experimental evidence that the quantum Hall liquid to insulator transition occurs at a universal ρ_{xx} value $\approx h/e^2$ [17], and since in our sample whenever $\rho_{xx} > h/e^2$, the derivative $d\rho_{xx}/dT < 0$, we take $\rho_{xx} = h/e^2$ as a threshold point for the IP. At each N , these threshold points have been converted to a critical ν for the onset of insulating behaviour and are plotted in figure 7 versus γ . The line connecting the points separates non-insulating from insulating states, and the resulting plot is therefore a phase diagram for the observed IP.

The first obvious feature of the phase diagram is the wrinkling in the phase boundary. This is caused by the re-entrance of the IP around several FQH states as discussed earlier. The limiting critical ν at low N is close to $1/5$, consistent with a low-disorder single-layer 2DES, while for the highest N where the system is effectively two layers in parallel, the IP is present for $\nu \lesssim 1/2$, i.e. $\nu \lesssim 1/4$ in each layer. This is reasonable considering that even at the largest N interlayer interactions are present in this system, as evidenced by the observation of a *correlated* $\nu = 1$ QHE at high N in the same well (see next section). Such interactions can move the WS ground state to $\nu \approx 1/4$ (for each layer), somewhat larger than $1/5$ expected if there are no interlayer interactions. We note that our measurements on wider quantum well samples indicate that in the high- N limit the onset of the IP indeed approaches $\nu \approx 2/5$, consistent with two high-quality parallel layers becoming insulating near $1/5$ filling in each layer in the absence of interlayer interactions. It appears that this regime is outside the density limits of our current sample.

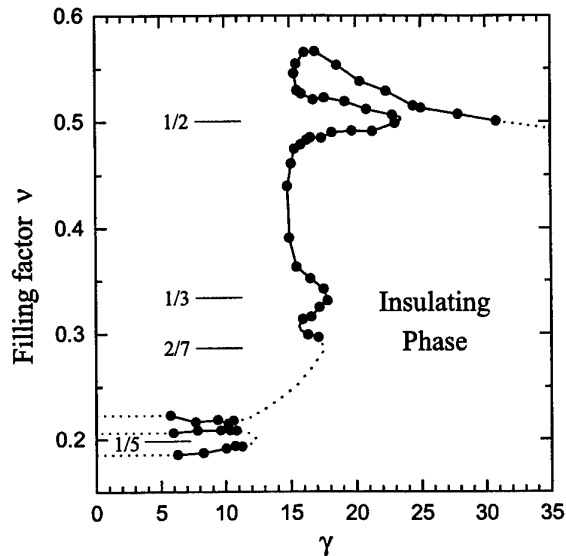


Figure 7. Phase diagram for the IP, constructed by plotting the critical ν for the onset of insulating behaviour versus the energy ratio γ . Dotted lines indicate regions outside our accessible field or density range.

We can examine in more detail the evolution of the IP as depicted in the phase diagram of figure 7 by making comparisons with figure 5. For intermediate N , as γ increases, the IP first remains close to $\nu = 1/5$ but then begins to move to higher ν in the range $12 < \gamma < 15$. This range is precisely bisected by $\gamma \approx 13.5$ (figure 5) where the 1C to 2C transition occurs. Then the IP moves

very quickly to $\nu \approx 1/2$ as evidenced by the nearly vertical phase boundary at $\gamma \approx 16$. As discussed earlier, and as evident from figure 5, this γ is centrally located in the parameter range necessary for the $1/2$ state.

Finally, we note that it is plausible that interlayer interactions can modify the ground state energies so that for appropriate parameters a crossing of the liquid and solid states occurs at the large fillings we identify (for example $\nu = 0.54$, i.e. $\nu > 1/4$ in each layer). Calculations [18] indicate that the effect of interlayer coupling can be particularly strong near the magnetoroton minimum and lead to the vanishing of the FQH liquid gap. This vanishing can be associated with an instability toward a ground state in which each of the layers condenses into a 2D WS [18].

In summary, our data clearly indicate that the IP we observe in this bilayer ES for $\gamma \gtrsim 13$ is a product of the same electron–electron interactions (both within the layers and between them) that give rise to the 1C to 2C FQH transition and the bilayer correlated $\nu = 1/2$ liquid state, evidence that the IP is indeed a 2C state with comparable inter- and intralayer correlations, and very likely a bilayer Wigner solid-like state.

5. Many-body, bilayer QHE at $\nu = 1$

In bilayer systems with appropriate parameters, the interlayer interactions can also lead to *correlated* QHE at *integral* fillings [19]. A particularly interesting example is the 2C Ψ_{111}^1 state at $\nu = 1$. In contrast to the 1C $\nu = 1$ QHE associated with Δ_{SAS} , the many-body, bilayer $\nu = 1$ incompressible state associated with Ψ_{111}^1 has been predicted to exhibit exotic properties such as neutral superfluid modes and a Kosterlitz–Thouless transition [20, 21], and has already revealed an unexpected in-plane B -driven transition (to another incompressible state) [22].

Data on ESs in wide wells [9] reveal that in these systems the ground state at $\nu = 1$ evolves continuously from a QH state stabilized by large Δ_{SAS} at low N to a many-body QH state stabilized by strong interlayer interaction at intermediate N . As N is further increased, we observe an incompressible-to-compressible transition. The unusual T and N dependence of the data as the transition boundary is approached is suggestive of an additional *finite-temperature* transition from a QHE to a compressible state, which is unique to bilayer systems. Here we briefly summarize these data.

Figure 3 presents Arrhenius plots of ρ_{xx} at $\nu = 1$ for several N in our 750 Å wide GaAs quantum well. The quasiparticle QHE excitation gaps Δ_1 determined from the slopes of the (low- T) activated regions of these plots, together with the measured and calculated Δ_{SAS} are shown in figure 8(a) as a function of γ . For comparison, the gaps for the $\nu = 1/2$ QH state in the same sample are also shown in figure 8(a). Several features of the data of figures 3 and 8 are noteworthy:

(i) While Δ_{SAS} decreases with increasing γ , Δ_1 increases and exceeds Δ_{SAS} by more than a factor of 3 at the highest measured γ .

(ii) For $N \gtrsim 10 \times 10^{10} \text{ cm}^{-2}$ ($\gamma \gtrsim 10$), the activated behaviour of R_{xx} versus $1/T$ starts rather abruptly below an

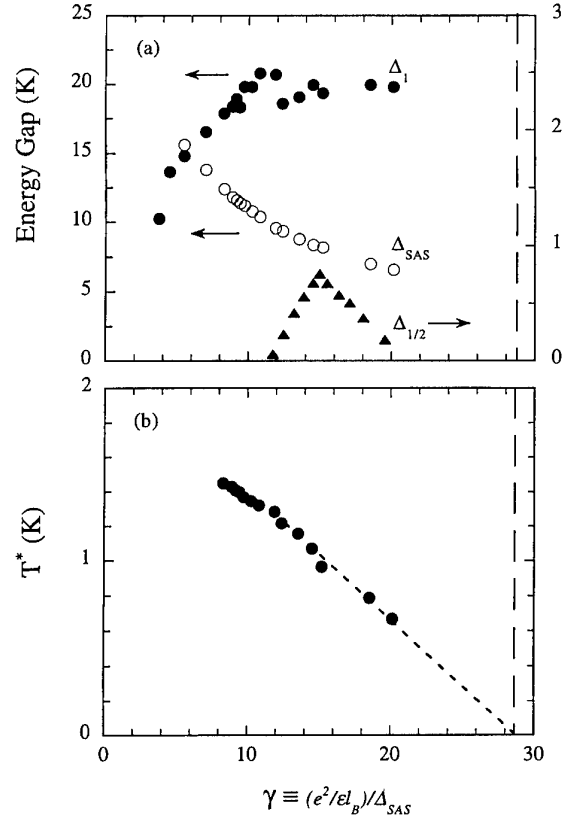


Figure 8. Plots of the measured $\nu = 1$ QHE energy gap (Δ_1) and T^* versus γ . The boundary for the collapse of the $\nu = 1$ QHE to a compressible state for this sample is shown by the vertical broken line. As indicated by the broken curve, the measured T^* extrapolates to zero at this boundary. For comparison, the measured Δ_{SAS} and the $\nu = 1/2$ QHE gap ($\Delta_{1/2}$) are also shown.

N -dependent temperature T^* . Above T^* , the R_{xx} minimum at $\nu = 1$ vanishes, i.e. R_{xx} becomes nearly independent of B and T . For $N \lesssim 10 \times 10^{10} \text{ cm}^{-2}$ ($\gamma \lesssim 10$), the Arrhenius plots show a smoother behaviour and Δ_1 gaps start to decrease with decreasing N .

(iii) The measured Δ_1 for $N > 10 \times 10^{10} \text{ cm}^{-2}$ ($\gamma > 12$) are approximately constant (≈ 20 K) and exceed T^* by more than an order of magnitude.

(iv) A plot of the measured T^* versus γ , presented in figure 8(b), shows that T^* decreases with increasing γ and extrapolates to zero at $\gamma \approx 29$. This γ corresponds to an incompressible-to-compressible phase boundary which we have observed in ESs in a number of wide quantum wells with varying width [9]: in these systems, for sufficiently large γ , we observe a collapse of the $\nu = 1$ QHE to a compressible state. In the present sample, we cannot reach this boundary because the required high N is not experimentally achievable but, based on our data on other samples, we expect the boundary to be at $\gamma \approx 29$ for this sample.

The above data demonstrate that the ground state of the ES in this wide quantum well at $\nu = 1$ evolves continuously from a 1C QH state at low N ($\gamma \lesssim 10$) stabilized by a large Δ_{SAS} to a 2C QH state at intermediate N ($10 \lesssim \gamma \lesssim 20$), and then makes a transition to a

metallic (compressible) state at large N ($\gamma > 29$). We believe that for intermediate N we are observing a bilayer QH state stabilized by comparable interlayer and intralayer correlations, possibly a 2C, Ψ_{111}^1 -like state [9]. Note that in this density range, γ for the $\nu = 1$ QH state overlaps with γ of the $\nu = 1/2$ QH state, a 2C liquid state which certainly requires interlayer correlations.

Finally, the data of figures 3 and 8 are collectively very unusual and qualitatively different from what is observed for the integral or fractional QHE in standard, single-layer 2DESs, or for the QHE at higher fillings such as $\nu = 3$ in the *same* wide quantum well sample [9]. The single-layer $\nu = 1$ QHE data typically exhibit a smoother saturation of the activated behaviour at high T and, as N is lowered, they show a larger R_{xx} (at any given T) and a smaller excitation gap [23]. This behaviour is very similar to what we observe for the $\nu = 3$ QHE in this sample [9] and for the $\nu = 1$ QHE at low N ($< 10 \times 10^{10} \text{ cm}^{-2}$) far away from the compressible boundary (figure 3). It is in sharp contrast to the $\nu = 1$ data at *high* N near the compressible boundary ($N > 10 \times 10^{10} \text{ cm}^{-2}$), where R_{xx} versus T^{-1} data appear to simply shift horizontally to lower T as N is raised.

While we do not have a clear understanding of these peculiar data, it is possible that a *finite-temperature* transition from a QHE to a compressible state is taking place, with T^* marking this transition [9]. Although unprecedented in a single-layer 2DES, finite- T transitions may occur in bilayer systems with appropriate parameters: examples include the Kosterlitz–Thouless transition theoretically proposed for the Ψ_{111}^1 state [20, 21], or a transition from a correlated (Ψ_{111}^1 -like) incompressible state to an uncorrelated, compressible state with $\nu = 1/2$ in each layer [24]. Details of such transitions and how they will quantitatively affect the transport properties are not known, however.

6. Summary

Magnetotransport data taken on an electron system in a wide quantum well with variable density reveal a striking evolution of its correlated states. While the data at low and high densities are consistent with single-layer and weakly coupled bilayer states respectively, at intermediate densities the data exhibit new QHE and insulating phases which are stabilized by both intralayer and *interlayer* electron correlations.

References

- [1] Willett R, Eisenstein J P, Stormer H L, Tsui D C, Gossard A C and English J H 1987 *Phys. Rev. Lett.* **59** 1776
Haldane F D M and Rezayi E H 1988 *Phys. Rev. Lett.* **60** 956
- [2] Clark R G, Haynes S R, Suckling A M, Mallett J R, Wright P A, Harris J J and Foxon C T 1989 *Phys. Rev. Lett.* **62** 1536
Eisenstein J P, Stormer H L, Pfeiffer L and West K W 1989 *Phys. Rev. Lett.* **62** 1540
- [3] Shayegan M, Jo J, Suen Y W, Santos M B and Goldman J 1990 *Phys. Rev. Lett.* **65** 2916
- [4] He S, Zhang F C, Xie X C and Das Sarma S 1990 *Phys. Rev. B* **42** 11 376
- [5] Suen Y W, Jo J, Santos M, Engel L W, Hwang S W and Shayegan M 1991 *Phys. Rev. B* **44** 5947
- [6] Suen Y W, Engel L W, Santos M B, Shayegan M and Tsui D C 1992 *Phys. Rev. Lett.* **68** 1379
- [7] Suen Y W, Santos M B and Shayegan M 1992 *Phys. Rev. Lett.* **69** 3551
- [8] Suen Y W, Manoharan H C, Ying X, Santos M B and Shayegan M 1994 *Phys. Rev. Lett.* **72** 3405
- [9] Lay T S, Suen Y W, Manoharan H C, Ying X, Santos M B and Shayegan M 1994 *Phys. Rev. B* **50** 17 725
- [10] Manoharan H C, Suen Y W, Santos M B and Shayegan M 1996 *Phys. Rev. Lett.* **77** 1813
- [11] Eisenstein J P, Boebinger G S, Pfeiffer L N, West K W and He S 1992 *Phys. Rev. Lett.* **68** 1383
- [12] See, for example, Sajoto T, Li Y P, Engel L W, Tsui D C and Shayegan M 1993 *Phys. Rev. Lett.* **70** 2321 and references therein
- [13] Halperin B I 1983 *Helv. Phys. Acta.* **56** 75
- [14] Yoshioka D, MacDonald A H and Girvin S M 1989 *Phys. Rev. B* **39** 1932
- [15] He S, Das Sarma S and Xie X C *Phys. Rev. B* **47** 4394
- [16] Sajoto T, Suen Y W, Engel L W, Santos M B and Shayegan M 1990 *Phys. Rev. B* **41** 8449
- [17] Shahar D, Tsui D C, Shayegan M, Bhatt R N and Cunningham J E 1995 *Phys. Rev. Lett.* **74** 4511
- [18] Oji H C, MacDonald A H and Girvin S M 1987 *Phys. Rev. Lett.* **58** 824
- [19] Chakraborty T and Pietilainen P 1987 *Phys. Rev. Lett.* **59** 2784
Fertig H A 1989 *Phys. Rev. B* **40** 1087
- [20] Wen X G and Zee A 1992 *Phys. Rev. Lett.* **69** 1811
Ezawa F and Iwazaki A 1992 *Int. J. Mod. Phys. B* **6** 3205
- [21] Moon K, Mori H, Yang K, Girvin S M, MacDonald A H, Zheng A H, Yoshioka D and Zhang S C 1995 *Phys. Rev. B* **51** 5138
- [22] Murphy S Q, Eisenstein J P, Boebinger G S, Pfeiffer L N and West K W 1994 *Phys. Rev. Lett.* **72** 728
- [23] See, for example, Usher A, Nicholas R J, Harris J J and Foxon C T 1990 *Phys. Rev. B* **41** 1129
- [24] Wen X G private communication

Breakdown of the integer quantum Hall effect at high currents in GaAs/AlGaAs heterostructures

S Kawaji

Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo 171, Japan

Abstract. Experimental results of the breakdown of the integer quantum Hall effect (QHE) measured at Gakushuin University for butterfly-type Hall bar samples fabricated from GaAs/AlGaAs heterostructure wafers with various electron concentrations and mobilities show that the onset of the breakdown of the QHE is caused by average current density or average Hall electric field. The result supports the idea that the edge channel picture is not a realistic description of the QHE. The critical breakdown Hall fields F_{cr} ($i = 2$ and 4) and F_{cr} ($i = 1$ and 3) are proportional to $B^{3/2}$ where B is the magnetic field at the centre of each quantized Hall plateau. The effective g -factor in this 2D system in strong magnetic fields is evaluated as $g_{eff} \approx 10$ from the experimental result F_{cr} ($i = 2$ and 4)/ F_{cr} ($i = 1$ and 3) ≈ 3 assuming F_{cr} ($i = 2$ and 4)/ F_{cr} ($i = 1$ and 3) = $\hbar\omega_c/g_{eff}\mu_B B$.

1. Introduction

In the integer quantum Hall effect (QHE), the Hall conductivity is quantized as $\sigma_{xy} = -ie^2/h$ for an integer i with zero diagonal conductivity $\sigma_{xx} = 0$ in degenerate two-dimensional electron systems (2DES) at low temperatures and in strong magnetic fields [1]. The QHE appears in a Hall bar sample as the Hall resistance is quantized as $R_H(i) = h/ie^2$ with zero diagonal resistance $R_{xx} = 0$. The quantum Hall effect was predicted by Ando *et al* [2] in 1975 by their theoretical study of the Hall effect in Landau levels in 2DES with attractive and repulsive short-range scatterers. In 1980 Kawaji and Wakabayashi [3] obtained experimental results by Hall current measurements for Si-MOSFETs which show that the Hall conductivity is quantized when the diagonal conductivity is zero as predicted by Ando *et al*. Their results of temperature dependence of the conductivities have shown that Anderson localization causes the quantum Hall effect as theoretically studied by Ando *et al*. In the same year von Klitzing *et al* [4] made a precision measurement of the Hall resistance $R_H = V_H/I$ and showed that $R_H(4) = h/4e^2$ with an accuracy of 3 ppm to the recommended value. Von Klitzing *et al* proposed that this is a new method for the determination of the fine structure constant $\alpha = \mu_0 ce^2/2h$ where μ_0 is the permittivity of a vacuum and c is the light velocity. This proposal opened a new era in the electrical resistance standards.

Breakdown of the QHE at high Hall fields was anticipated by early Corbino disc experiments for Si-MOSFETs where the finite gate voltage regions for $\sigma_{xx} = 0$ at low source-drain fields disappear when the source-drain field increases [1,5]. In the early 1980s several

experimental and theoretical studies of the breakdown of the QHE were reported [6]. Experimental researches into the breakdown were mainly carried out for Hall bar samples because the main concern in standards laboratories was a high current in a Hall bar which generates a high Hall voltage for high-precision measurements in resistance standards.

Recently, the breakdown of the QHE at high currents has been a subject of renewed interest as a tool to find current or field distributions in the QHE state. In the last decade, the quantized Hall resistance (QHR) and zero longitudinal resistance in the QHE has been discussed as a non-local property of a 2DES in strong magnetic fields based on an edge channel picture [7]. This picture is different from the Anderson localization picture which describes the QHE as a local property of a 2DES. Breakdown experiments have been expected to clarify the mechanism of the QHE.

This paper briefly describes the results of a series of experiments by our group on the breakdown of the QHE.

2. Samples

The QHE state is characterized by non-dissipative electric current whose direction is orthogonal to the direction of the electric field, i.e. a Hall electric field, where the diagonal resistance in the Hall bar is zero, i.e. $R_{xx} = 0$. Such a non-dissipative electronic state only appears near the central part of the Hall bar sample. Joule heat is always generated at both ends of the sample, i.e. in transition regions between the 2DES and two current electrodes, the source electrode and the drain electrode [1]. The breakdown of the QHE which we want to measure is the appearance of dissipation

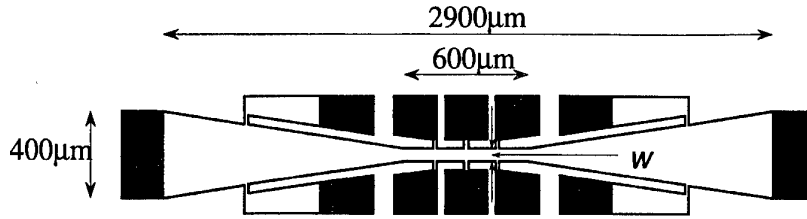


Figure 1. Structure of butterfly-type I samples [9, 10].

in a part of a Hall bar other than the parts near the current electrodes. In order to do this, we have to eliminate effects of the source and the drain electrodes. We know that in high-precision measurements of the QHR in Si-MOSFETs the resistances at the source and the drain electrodes appear in narrow Hall bar samples [8]. The Si-MOSFET samples with wide source and drain electrodes (~ 2 mm) do not show electrode resistance at a current level of $10 \mu\text{A}$.

In a series of our experiments we fabricated specially designed Hall bar samples (butterfly-type Hall bar) from GaAs/Al_{0.3}Ga_{0.7}As heterostructure wafers by photolithography and wet chemical etching. The typical electrode structure, called the butterfly-type I, is shown in figure 1 [9,10]. In order to keep the same energy dissipation condition at the source and drain electrodes in different samples, all the samples have the same source and drain electrode width of $W = 400 \mu\text{m}$. Moreover, in order to equilibrate the state of electrons in the central part of the sample where the measurement is carried out, all samples have a large total length of $L = 2900 \mu\text{m}$. The central part is $600 \mu\text{m}$ long and has a different width w between 10 and $120 \mu\text{m}$. The length $\ell = 600 \mu\text{m}$ is found in three pairs of potential probes. The width of each sample is linearly narrowed from both current electrodes to the ends of the central part as shown in figure 1. In another series of experiments we used samples (butterfly-type II) with $W = 400 \mu\text{m}$, $L = 2600 \mu\text{m}$ and $\ell = 120 \mu\text{m} + 12w$ [11]. The length of $120 \mu\text{m}$ is found in three pairs of potential probes. The width w is fabricated as 3, 4, 5, 10, 20, 50 and $60 \mu\text{m}$.

We also fabricated conventional rectangular Hall bar samples with a length between the source and the drain electrodes of $L = 600 \mu\text{m}$ and three pairs of potential probes. This structure is same as the central part of the butterfly-type I [12].

3. Experimental results

Measurements were made mostly at 0.5 K. The quantized Hall plateau region was observed in a trace of magnetic field dependence of R_{xx} . Typical traces are shown in figure 2 [11]. The critical current I_{cr} was determined from the current at which the plateau disappears, i.e. the diagonal resistance $> 0.1 \Omega$. The rise in the diagonal resistance at the breakdown is so abrupt that the critical current can be determined without any uncertainty affecting the results. An example of the sample width dependence of the critical current is shown in figure 3 [11]. The width of the central

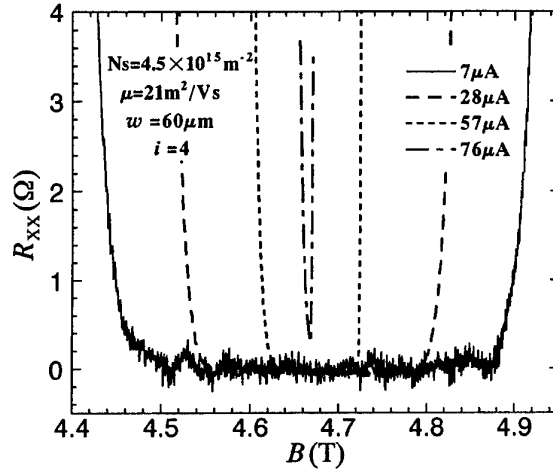


Figure 2. Diagonal resistance $R_{xx}(i = 4)$ versus change in the magnetic field B [11].

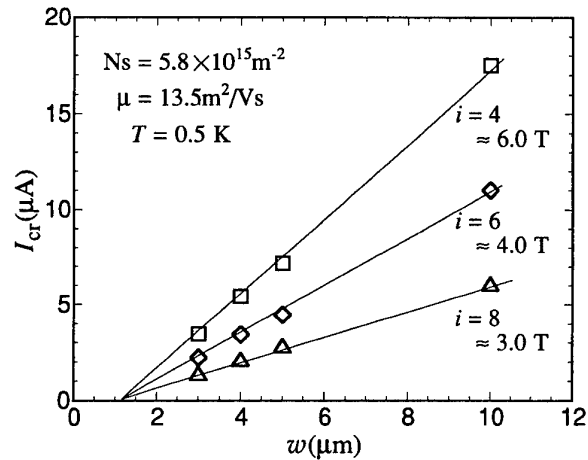


Figure 3. Critical current of the breakdown I_{cr} versus sample width w [11].

part is reduced from w to the effective sample width w' by chemical etching and formation of space charge layers. As shown in figure 3, the critical current is proportional to the effective sample width. Therefore, we used the effective width for calculation of average critical current densities J_{cr} , or average critical Hall electric fields F_{cr} . We used six wafers in which electron concentrations N_s and mobilities

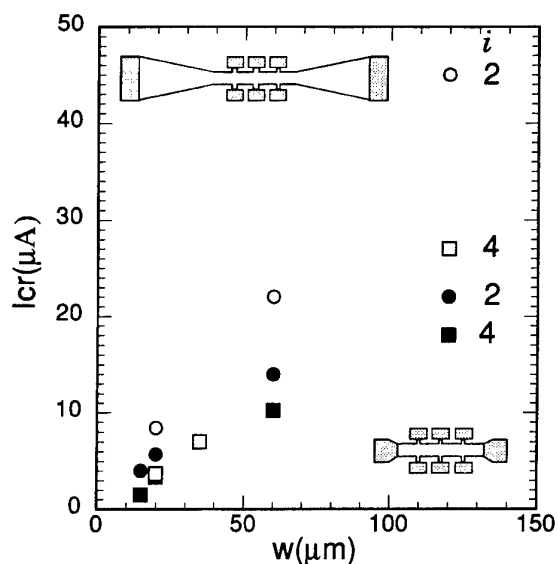


Figure 4. Critical current of the breakdown I_{cr} versus sample width w in butterfly-type I samples (open symbols) and rectangular samples (full symbols) [12].

μ ranged from $1.4 \times 10^{15} \text{ m}^{-2}$ to $5.8 \times 10^{15} \text{ m}^{-2}$ and from $13.5 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ to $27 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$.

We compared breakdown features between the butterfly-type I samples and conventional rectangular samples fabricated on the same heterostructure wafer which has $N_s = 1.6 \times 10^{15} \text{ m}^{-2}$ and $\mu = 35 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. An example of the width dependence of the critical breakdown current is shown in figure 4 [12]. The conventional rectangular samples show a sublinear dependence of I_{cr} on w , while the butterfly-type I samples show a linear dependence.

We have extended breakdown measurements for higher-mobility butterfly-type samples. Butterfly-type II samples fabricated from a wafer which have $N_s = 3.5 \times 10^{15} \text{ m}^{-2}$ and $\mu = 75 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ showed the same linear dependence on w as those measured in low-mobility samples. A butterfly-type I sample fabricated from a wafer of $N_s = 3.0 \times 10^{15} \text{ m}^{-2}$ and $\mu = 110 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ showed almost the same F_{cr} as those of low-mobility samples [13].

Experimental results for critical breakdown electric fields F_{cr} in the Hall plateaus with the plateau quantum numbers $i = 1, 2, 3, 4$ are plotted against the magnetic field at the centre of each plateau in figure 5 [13]. Figure 5 shows that F_{cr} ($i = 2, 4$) lie on a single straight line and F_{cr} ($i = 1, 3$) lie on another parallel single line in the log-log plot and their slopes are $3/2$.

4. Discussion

An important result of our experiments carried out by using butterfly-type Hall bars is that the onset of the breakdown of the QHE is not caused by the electric current but the average current density or the average Hall electric field. This characteristic feature does not depend on electron mobility which ranges between 13.5 and $110 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. The result strongly supports a homogeneous current distribution

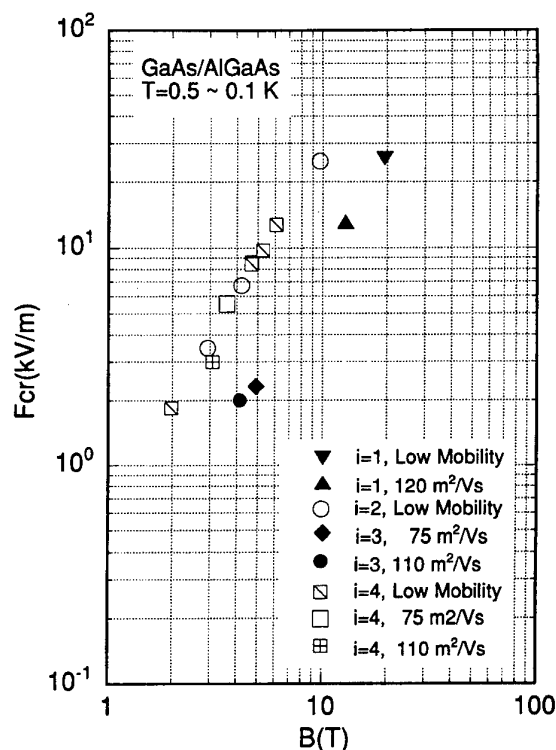


Figure 5. Critical breakdown field F_{cr} versus magnetic field B [13]. Low mobility: $13.5 \leq \mu \text{ (m}^2 \text{ V}^{-1} \text{ s}^{-1}) \leq 27$.

in a Hall bar, at least at high currents such as those used in resistance standards. In other words, the edge state picture is not a realistic description of the QHE. The edge channel picture might be able to describe phenomena in Hall bars at small currents $I < 0.1 \mu\text{A}$ [14]. The phenomena which can be observed at such small currents, however, should not be called the QHE because we cannot use the Hall resistance at a current level of $0.1 \mu\text{A}$ as resistance standards. We note here that a quantized Hall conductivity can be observed without an edge by measurements using microwave frequency [15] and by measurements for Corbino disc samples [16].

Recently Balaban *et al* [17] claimed that their critical current I_{cr} measured in high-mobility samples depends sublinearly on the width w . The electron mobilities of their samples are 70 and $90 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. They also claimed that the low-mobility sample with $\mu = 12 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ showed linear dependence of I_{cr} on w but the dependence changed into sublinear and I_{cr} decreases after long illumination where the mobility changed to $21 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ [18]. They attributed the linear dependence of I_{cr} on w to inhomogeneity of the samples. However, our results for F_{cr} in high-mobility butterfly-type samples with $\mu = 75 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ and $\mu = 110 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ showed almost same results as observed in low-mobility samples. Moreover, as shown in figure 4, conventional rectangular Hall bar samples showed smaller I_{cr} than those measured in the butterfly-type samples and their dependence on w is sublinear. Based on these results we consider that the sublinear dependence of I_{cr} on w , which Balaban *et al*

observed [17,18] probably arises from effects of current electrodes in short samples, though, unfortunately, Balaban *et al* did not describe the lengths of their samples.

We summarized in figure 6 F_{cr} ($i = 2, 4$) results reported so far by other authors [19–26] and our results. Our F_{cr} data are no smaller than other authors' results measured by using normal Hall bars and Corbino discs. This shows that our F_{cr} data are probably not affected by any effects of current electrodes. In figure 6, Stormer *et al*'s results (open circles 4) for $i = 2$ and $i = 4$ were measured by a Corbino disc [22]. Therefore they do not contain the effects of current electrodes expected in Hall bar samples. Their values are, however, smaller than many other results measured by Hall bar samples and the magnetic field dependence is much weaker than our results. Their Corbino disc results might contain other uncontrollable sample parameters, such as inhomogeneity in the wafer, than simple electrode effects. We note here that Blik *et al*'s [27] result measured for samples with a narrow constriction is one order of magnitude larger than our results.

The magnetic field dependence of F_{cr} is a key which opens the door to a possible mechanism for the breakdown. The mechanism of abrupt phonon emission analogous to Cerenkov radiation, proposed by Streda and von Klitzing [28] and Stormer *et al* [22], and an electron-heating mechanism such as discussed by Komiyama *et al* [23] lead to a dependence of $F_{cr} \propto B$. Our result of $F_{cr} \propto B^{3/2}$ suggests that a possible mechanism for the breakdown is the production of dissipative carriers by inter-Landau level tunnelling of electrons due to small scatterers from the highest filled Landau level to the next empty Landau level. The spatial extent of the electron wavefunction of a Landau level along the Hall electric field is proportional to $B^{-1/2}$ and is described by a harmonic oscillator wavefunction [1]. The energy splitting between adjacent Landau levels is proportional to B irrespective of their origin, i.e. Landau gaps or spin gaps. Therefore, the critical electric field at the onset of electron tunnelling is expected to be proportional to $B^{3/2}$ irrespective of tunnelling processes, i.e. a process due to a single transition or a process due to higher-order multiple transitions.

Eaves and Sheard (E-S) [29] have derived a simple expression for the critical breakdown field based on the inter-Landau level single process tunnelling as

$$F_{cr}(E-S) = \frac{\hbar\omega_c}{e\ell_B[(2n+1)^{1/2} + (2n+3)^{1/2}]} \quad (1)$$

where $\ell_B = (\hbar/eB)^{1/2}$ is the magnetic length and n is the Landau level quantum number, i.e. $n = 0$ for $i = 1$ and 2 and $n = 1$ for $i = 3$ and 4 . This equation leads to a relation $F_{cr}(i) \propto B^{3/2}$. Here, $\ell_B(2n+1)^{1/2}$ is the spatial extent of the wavefunction of an electron in the highest filled Landau level and $\ell_B(2n+3)^{1/2}$ is that of an electron in the lowest empty Landau level. Thus the product of the denominator of equation (1) and the Hall electric field gives the potential energy difference of an electron along the distance where its wavefunction and the wavefunction of a state in the next higher Landau level overlap each other. However, our experimental results are different from the

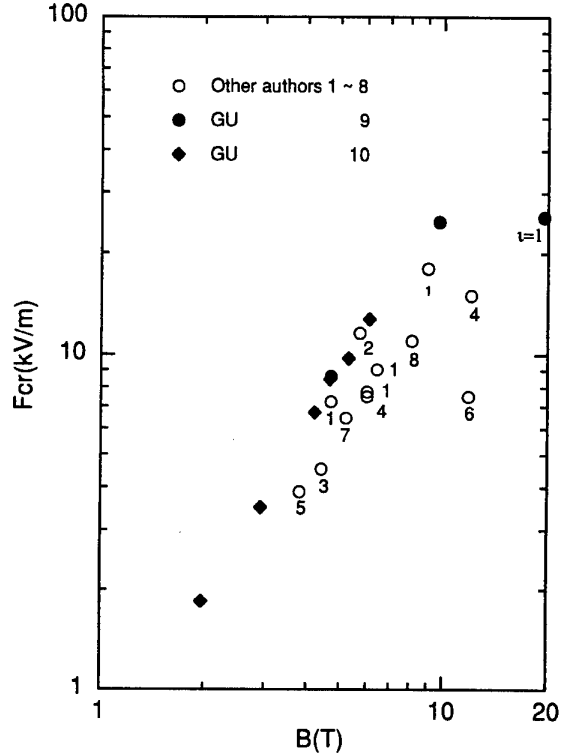


Figure 6. Critical breakdown field $F_{cr}(i = 2, 4)$ versus magnetic field B reported by other authors (open circles) and by the GU group (full symbols). 1, Ebert *et al* [19]; 2, Cage *et al* [20]; 3, Kuchar *et al* [21]; 4, Stormer *et al* [22]; 5, Komiyama *et al* [23]; 6, Cage *et al* [24]; 7, Boisen *et al* [25]; 8, Ahler *et al* [26]; 9, Kawaji *et al* [10, 11]; 10, Okuno *et al* [11]. Sample specifications of GU group. GU9: $N_s(10^{15} \text{ m}^{-2}) = 4.7$, $\mu(\text{m}^2 \text{ V}^{-1} \text{ s}^{-1}) = 21$, GU10: $5.8 \geq N_s(10^{15} \text{ m}^2) \geq 1.4$, $13.5 \leq \mu(\text{m}^2 \text{ V}^{-1} \text{ s}^{-1}) \leq 27$.

prediction of equation (1) in their magnitude and in their dependence on the Landau level quantum number. Eaves and Sheard's expression leads to a result which is more than one order of magnitude larger than our experiments, i.e. $[F_{cr}(E-S)/F_{cr}(\text{Expt.})]_{i=4} \approx 20$. According to their expression, $F_{cr}(i = 2)$ should be larger than $F_{cr}(i = 4)$, i.e. $[F_{cr}(E-S)_{i=2}/F_{cr}(E-S)_{i=4}] = 1.5$, while they are the same in our experimental results.

When we take higher-order tunnelling processes into account, in which an electron can tunnel through multiple paths via higher and lower energy states, F_{cr} may be reduced from the value calculated by equation (1). Recently Kawabata [30] calculated the critical field due to higher-order tunnelling and obtained a result $F_{cr}(\text{Kawabata})/F_{cr}(E-S) \sim 1/3$. The relation $F_{cr} \propto B^{3/2}$ is held for higher-order tunnelling. The higher-order tunnelling may also eliminate the dependence of the Landau level quantum number in the inter-Landau level transition [30]. The enhancement of the inter-Landau level transition by spatially extended scatterers proposed by Trugman [31] reduces the critical field but their result leads to a weaker magnetic field dependence of $F_{cr} \propto B$. Quite recently Ishikawa *et al* calculated the inter-Landau level tunnelling caused by Landau level broadening due to a Hall electric

field [32]. Ishikawa *et al*'s representation of 2DES in a strong magnetic field is based on von Neuman's magnetic lattice [33]. They calculated the level broadening due to the electric field F and obtained that the level width is given by $\Gamma = (2\pi)^{1/2} e \ell_B F$. They consider that the electron tunnelling occurs at the field when the Landau level width is equal to the energy splitting. Then the critical field is given by $F_{cr}(\text{Ishikawa}) = \hbar\omega_c / (2\pi)^{1/2} e \ell_B$ which is independent of the plateau quantum number i and $F_{cr}(\text{Ishikawa})/F_{cr}(\text{E-S}) \sim 1$.

Our results in figure 5 show that the ratio $F_{cr}(i = 2, 4)/F_{cr}(i = 1, 3)$ is about 3 at a given magnetic field. Strong spin-orbit interaction in GaAs allows spin-flip tunnelling from the filled lower spin-split spin polarized Landau level to the upper spin-split Landau level. Therefore, we consider this ratio to be equal to the ratio of the Landau level splitting and the Zeeman splitting $\hbar\omega_c/g_{\text{eff}}\mu_B B$ where g_{eff} is the exchange-enhanced effective g -factor of an electron and μ_B is the Bohr magneton. The relation $F_{cr}(i = 2, 4)/F_{cr}(i = 1, 3) = \hbar\omega_c/g_{\text{eff}}\mu_B B$ is easily given when the inter-Landau level tunnelling mechanism holds in the breakdown of the QHE in the Landau gaps and in the Zeeman gaps. Even if this breakdown mechanism does not hold, the relation given above between the critical fields and the energy splittings probably holds though it is hard to explain the dependence of $F_{cr} \propto B^{3/2}$ by other mechanisms. Then we can determine the effective g -factor of an electron in the present system. Spin-splitting can be ignored in the breakdown in the plateaus with $i = 2$ and 4 because the g -factor of an electron in GaAs is very small, i.e. $g = -0.44$ [34]. However, the g -factor should be enhanced by the exchange effect observed for quantized plateaus with $i = 1$ and 3. From the result $F_{cr}(i = 2, 4)/F_{cr}(i = 1, 3) \approx 3$ in figure 5, we have a result $g_{\text{eff}} \approx 10$. This result is consistent with Usher *et al*'s $g_{\text{eff}} = 7.3$ [35]. They obtained the same magnetic field dependence of the activation energies of R_{xx} in both energy gaps for $i = 1$ and $i = 3$ and derived the above result. Our result is also consistent with a recent result in experiments on skyrmions by Barrett *et al* [36] who discussed the enhancement of Zeeman splitting by a factor of ~ 10 to explain temperature dependence of the Knight shift in optically pumped NMR measurements near the Landau level filling factor $\nu = 1$.

In conclusion we have carried out measurements of the critical current at the onset of the breakdown of the QHE in butterfly-type Hall bar samples made from GaAs/Al_{0.3}Ga_{0.7}As heterostructure wafers which have different electron concentrations and mobilities. We have observed that the onset of the breakdown is caused not by the current but by the average current density or the average Hall electric field and the critical field is proportional to $B^{3/2}$ where B is the magnetic field at the plateau centre. The former shows that the current distribution is uniform at a high current level such as those used in resistance standards. The latter suggests that the inter-Landau level tunnelling due to small scatterers is a possible breakdown mechanism. The tunnelling is probably not a simple process but a higher-order multiple process. The effective g -factor of an electron in the present system in strong magnetic

fields is evaluated to be ≈ 10 from the ratio of the critical field for the plateau quantum number $i = 2$ and 4 and that for $i = 1$ and 3 assuming $F_{cr}(i = 2 \text{ and } 4)/F_{cr}(i = 1 \text{ and } 3) = \hbar\omega_c/g_{\text{eff}}\mu_B B$.

Acknowledgments

The author would like to thank colleagues in Gakushuin University, in particular A Kawabata for his helpful discussions, and T Fukase and T Goto in IMR, Tohoku University, for their help in high-magnetic-field experiments. The author is obliged to K Ishikawa for his preprint and discussions. This work is supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture, Japan.

References

- [1] Kawaji S 1994 *Proc. Int. Symp. on the Foundation of Quantum Mechanics (Tokyo, 1983)* ed S Kamefuchi *et al* (Tokyo: Physical Society of Japan) p 327
- [2] Ando T, Matsumoto Y and Uemura Y 1975 *J. Phys. Soc. Japan* **39** 279
- [3] Kawaji S and Wakabayashi J 1981 *Proc. Int. Conf. on Physics in High Magnetic Fields (Hakone, 1980)* ed S Chikazumi and N Miura (Berlin: Springer) p 284
- [4] von Klitzing K, Dorda G and Pepper M 1980 *Phys. Rev. Lett.* **45** 449
- [5] Kawaji S and Wakabayashi J 1976 *Surf. Sci.* **58** 254
- [6] See references in [10]
- [7] Buttikker M 1988 *Phys. Rev. B* **38** 9375
- [8] Yoshihiro K, Kinoshita J, Inagaki K, Yamanouchi C, Endo T, Murayama Y, Koyanagi M, Yagi A, Wakabayashi J and Kawaji S 1986 *Phys. Rev. B* **33** 6874
- [9] Kawaji S, Hirakawa K and Nagata N 1993 *Physica B* **184** 17
- [10] Kawaji S, Hirakawa K, Nagata N, Okamoto T, Goto T and Fukase T 1994 *J. Phys. Soc. Japan* **63** 2303
- [11] Okuno T, Kawaji S, Ohnri T, Okamoto T, Kurata Y and Sakai J 1995 *J. Phys. Soc. Japan* **64** 1881
- [12] Fukano A, Kawaji S, Okamoto T, Kurata Y and Sakai J unpublished
- [13] Kawaji S, Fukano A, Okamoto T, Goto T and Fukase T unpublished
- [14] von Klitzing K 1995 *Physica B* **116** 111
- [15] Kuchar F, Meisel R, Weimann G and Schlapp W 1986 *Phys. Rev. B* **33** 2965
- [16] Dolgoplov V T, Zhitenev N B and Shashkin A A 1990 *Pis'ma Zh. Eksp. Teor. Fiz* **52** 826; 1990 *JETP Lett.* **52** 196
- [17] Balaban N Q, Meirav U, Strikman H and Levine Y 1993 *Phys. Rev. Lett.* **71** 1443
- [18] Balaban N Q, Meirav U and Strikman H 1995 *Phys. Rev. B* **52** R5503
- [19] Ebert G, von Klitzing K, Ploog K and Weiman G 1983 *J. Phys. C: Solid State Phys.* **16** 5441
- [20] Cage M E, Dziuba R F, Field B F, Williams E R, Girvin S M, Gossard A C, Tsui D C and Wagner R J 1983 *Phys. Rev. Lett.* **51** 1347
- [21] Kuchar F, Bauer G, Weiman G and Burkhard H 1984 *Surf. Sci.* **142** 196
- [22] Stormer H L, Chan A M, Tsui D C and Hwang J C M 1985 *Proc. 17th Int. Conf. on the Physics of Semiconductors (San Francisco, 1984)* ed D C Chadi and W A Harrison (Berlin: Springer) p 267
- [23] Komiyama S, Takamasu T, Hiyamizu S and Sasa S 1985 *Solid State Commun.* **54** 479
- [24] Cage M E, Marullo Reedtz G, Yu D Y and Vandegrift C T 1990 *Semicond. Sci. Technol.* **5** 351

- [25] Boisen A, Boggild P, Kristensen A and Lindelof P E 1994 *Phys. Rev. B* **50** 1957
- [26] Ahler F J, Hein G, Scherer H, Blik L, Nickel H, Rosch R and Schlapp W 1993 *Semicond. Sci. Technol.* **8** 2062
- [27] Blik L, Braun E, Hein G, Lose V, Niemeyer J, Weiman G and Schlapp W 1986 *Semicond. Sci. Technol.* **1** 110
- [28] Streda P and von Klitzing K 1984 *J. Phys. C: Solid State Phys.* **17** L483
- [29] Eaves L and Sheard F W 1986 *Semicond. Sci. Technol.* **1** 346
- [30] Kawabata A private communication
- [31] Trugman S A 1989 *Phys. Rev. Lett.* **62** 579
- [32] Ishikawa K, Maeda N and Tadaki K *Phys. Rev. B* submitted
- [33] Ishikawa K, Maeda N and Tadaki K 1995 *Phys. Rev. B* **51** 5048
- [34] Weisbuch C and Hermann C 1977 *Phys. Rev. B* **15** 816
- [35] Usher A, Nickolas R J, Harris J J and Foxon C T 1990 *Phys. Rev. B* **41** 1129.
- [36] Barrett S E, Dabbagh G, Pfeiffer L N, West K W and Tycko R 1995 *Phys. Rev. Lett.* **71** 1443

Nanolithography

D K Ferry, M Khoury, D P Pivin Jr, K M Connolly, T K Whidden,
M N Kozicki and D R Allee

Department of Electrical Engineering, Arizona State University, Tempe,
AZ 85287-5706, USA

Abstract. Nanolithography is the technology of fabricating sub-0.1 μm structures for use in either electronic devices or novel experiments in physics. While this has predominantly been the province of electron-beam lithography, new approaches explore the use of scanning probe techniques. Here, we discuss the limitations of electron-beam lithography, the development of novel SiO_2 resists, and some uses of scanning-probe lithography.

1. Introduction

Since the introduction of integrated circuits some four decades ago, the density of individual transistors on a single circuit chip has doubled approximately every three years, and it is expected that minimum linewidths will approach 0.07 μm within the next decade. This growth has fuelled a continuing improvement in the tools that are used for photolithographic exposure and fabrication of these circuits, and it is expected that optical lithography will continue to be the methodology of choice down to dimensions of the order of 0.12–0.15 μm . Beyond that, new approaches will be necessary. On the other hand, the fabrication of research devices at much smaller dimensions has always utilized more advanced nanolithographic techniques, usually electron-beam (e-beam) lithography [1–4]. The utility of direct writing with the e-beam system also has been a key to the fabrication of specialized structures in which to investigate novel physical effects, particularly low-dimensional structures [5].

Many approaches have been used to define novel sub-0.1 μm structures by nanolithography [6]. The highest resolution of these approaches remains electron-beam lithography and its newer offspring, scanning-probe lithography, in which a scanning tunnelling microscope (STM) or an atomic force microscope is used to create the exposure [7–10]. In this paper, we will review the limits of electron-beam lithography using both the standard photoresist polymethylmethacrylate (PMMA) and a novel new resist—silicon dioxide—that is both fast and exhibits high resolution when processed with the proper developers. In addition, we will look at some novel work on the use of the scanning probe microscope for oxidation lithography.

2. Nanolithography

The lithographic process, as it is used to create low-dimensional systems, is really composed of three steps. These are: (i) exposure of a resist material, (ii) development of the resist material, and (iii) pattern transfer. One cannot address any of these three steps in isolation, since, for

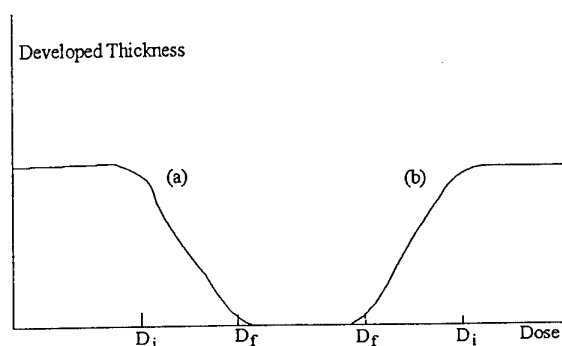


Figure 1. The log exposure versus dose curves for a positive (a) and negative (b) resist. The contrast is the maximum slope of the transition region as defined by equation (1).

example, the resolution that one can attain in a given resist depends not only upon the inherent resolution of the resist material (polymer size, etc), but also on the actual exposure that is used and the development procedure and chemistry that is used to fix the pattern into the resist. Finally, pattern transfer depends upon the robustness of the resist to the process chemistry involved in the transfer itself, which may vary over the very different processes ranging from reactive-ion etching to metal deposition and lift-off.

In defining the resolution of a resist, the most common attribute is the *contrast* γ , which is defined as

$$\gamma = |\ln(D_f/D_i)|^{-1}. \quad (1)$$

Here, for a positive (negative) resist, D_i is the maximum (minimum) dose for which the developed resist thickness is unchanged and D_f is the minimum (maximum) dose for which the developer can completely remove the resist. The magnitude sign arises since for negative resists $D_f < D_i$, while for positive resists $D_f > D_i$. That is, in a negative resist, the unexposed regions are removed, while in a positive resist the exposed regions are removed by the developer. This difference is indicated in figure 1. Since

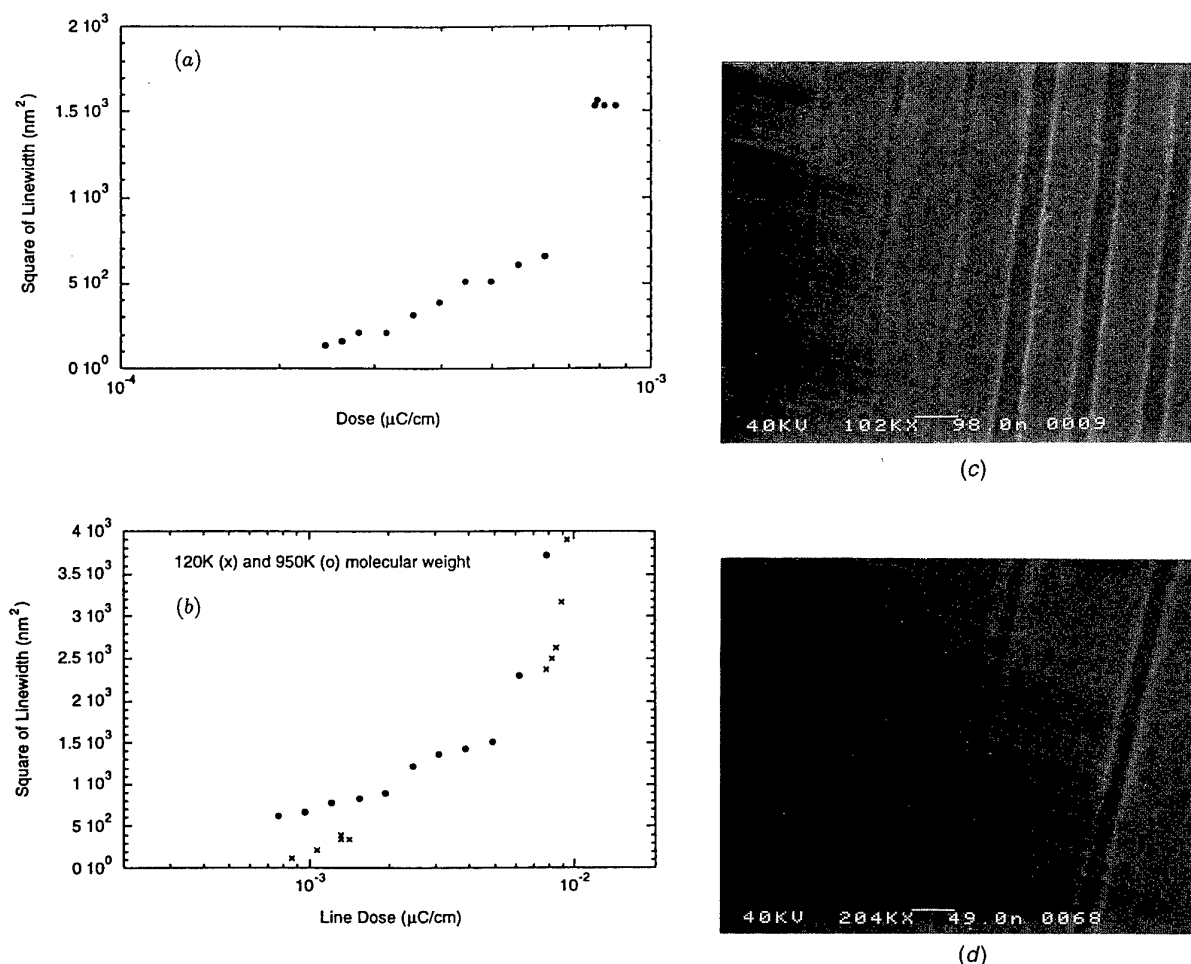


Figure 2. The square of the linewidth obtained with electron-beam exposure of PMMA as a function of the dose for (a) 15 K molecular weight, and (b) 120 K and 950 K molecular weights. The intercept on the horizontal axis gives the critical dose, while the slope of the linear part of the curves yields the value for σ discussed in the text.

most resists do not follow an ideal curve, γ is usually taken from the maximum slope of the dose versus exposure curve. Most high-resolution e-beam lithography for low-dimensional systems has utilized PMMA as the resist of choice. In our group, we have achieved very high resolution in PMMA by the development of a unique combination of the standard developers methyl-isobutyl ketone, cellosolve and methyl-ethyl ketone [11], which produces $\gamma > 11$.

3. The limits of PMMA

The polymeric material PMMA (or Plexiglass as it is sometimes known) has been used for many years as the high-resolution resist of choice in high-resolution electron-beam lithography [12]. As mentioned, this has allowed the fabrication of ultrasmall transistors with gate lengths in the 20–40 nm range [1–4, 13]. However, it is generally felt that one can go to even smaller sizes. The intrinsic limit in PMMA has generally been thought to be about 10 nm [14, 15], although recent work has suggested that lines as small as 6 nm are possible [16, 17]. During e-beam irradiation, the long chains of PMMA are fragmented into smaller chains, which are later dissolved in a solvent

(developer). As discussed above, there are several factors in evaluating a high-resolution resist, one of which is the properties of the resist and developer. A second factor in the resolution achieved is the properties of the exposing beam itself. Here, beam diameter is not the limiting factor. Rather, the electron interaction with the solid, for example the forward and back-scattered electrons, contributes to the exposure. Since most organic compounds are affected by electrons with energies in the 5 eV range, the effects of the secondary electrons have to be carefully appraised.

To determine the distribution of electrons, several experimental studies have been conducted, usually in the form of resolution tests [14]. We have also carried out these studies in order to probe the limiting resolution of PMMA. The normal beam cross section produces a Gaussian exposure pattern, which may be written as

$$P(x) \sim \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (2)$$

where the standard deviation σ describes the exposure latitude. Broers [14] found values in the range 7.8–9.8 nm (corrected for the factor of 2 he omitted from equation (2)), and concluded that the minimum linewidth would be limited

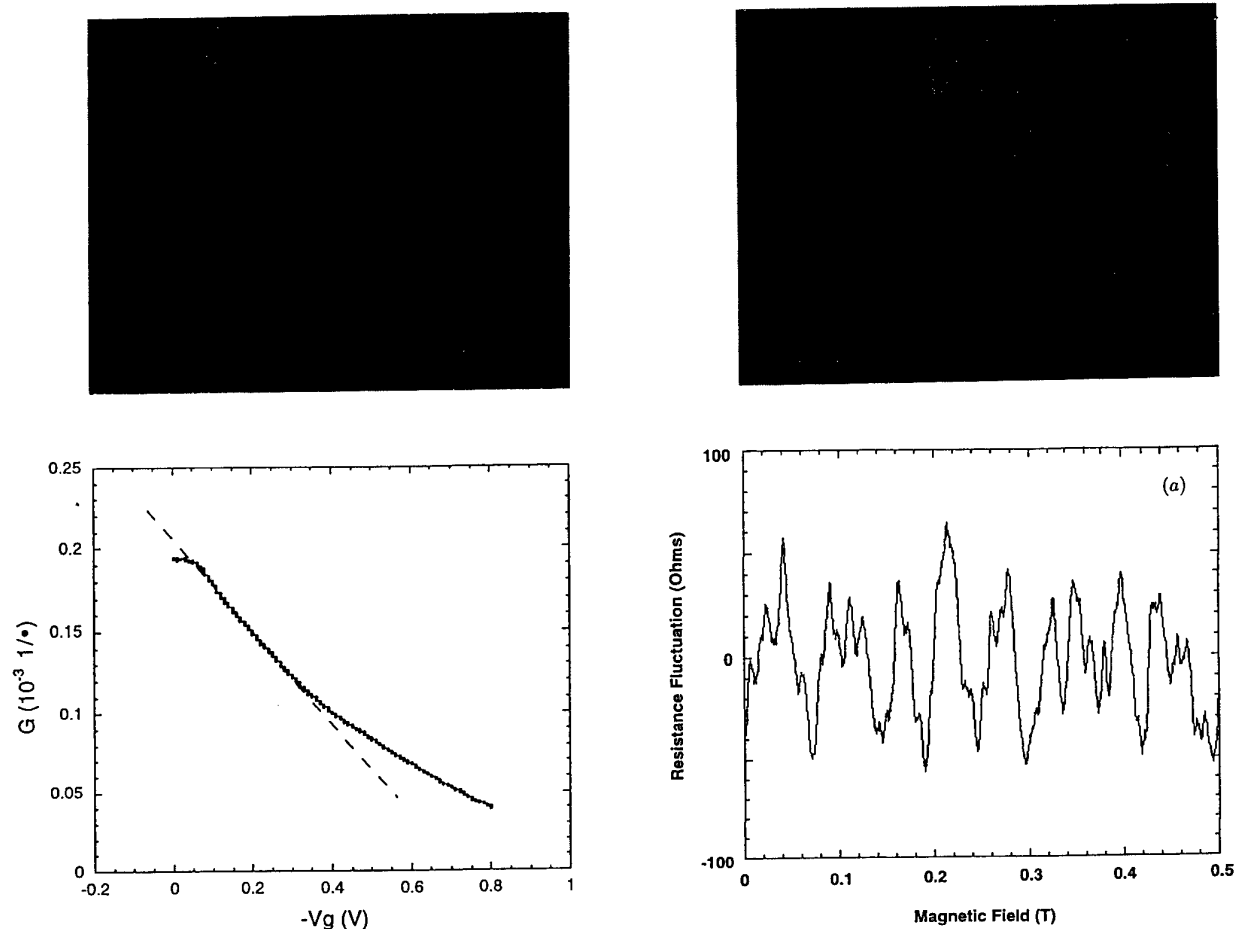


Figure 3. Resistance of an in-plane gated quantum wire produced by subsequent pattern transfer during RIE. The top part shows a picture of the wire itself, which has been defined by the trench etched into the heterostructure.

to the value of this parameter. In figure 2, we show the square of the linewidth obtained with 40 kV exposure, as a function of the actual dose used, for molecular weights varying from 15 K to 950 K [17]. Usually, the heavier molecular weight is used in the belief that it produces better resolution. Values of σ for all the molecular weights were found to be approximately 8 nm, although minimum linewidths varied from about 15 nm in 15 K molecular weight to 6–7 nm in 950 K molecular weight.

One important aspect of lithography is the resultant pattern transfer that must be done following exposure and development of the resist. In most cases, this is done by lift-off of metal, and we have used such procedures to fabricate ultrasmall gate length field-effect transistors [1] as well as other quantum structures such as surface superlattices of quantum dot arrays [13,18]. Another procedure is to use the mask as an etch mask for subsequent reactive-ion etching (RIE). In the top part of figure 3, an in-plane-gated quantum wire produced by this method is pictured, in which thin trenches have been etched into the GaAs/AlGaAs heterostructure. The advantage of in-plane gates is a linear variation of wire width with gate voltage, rather than square-root variation as usually obtained with

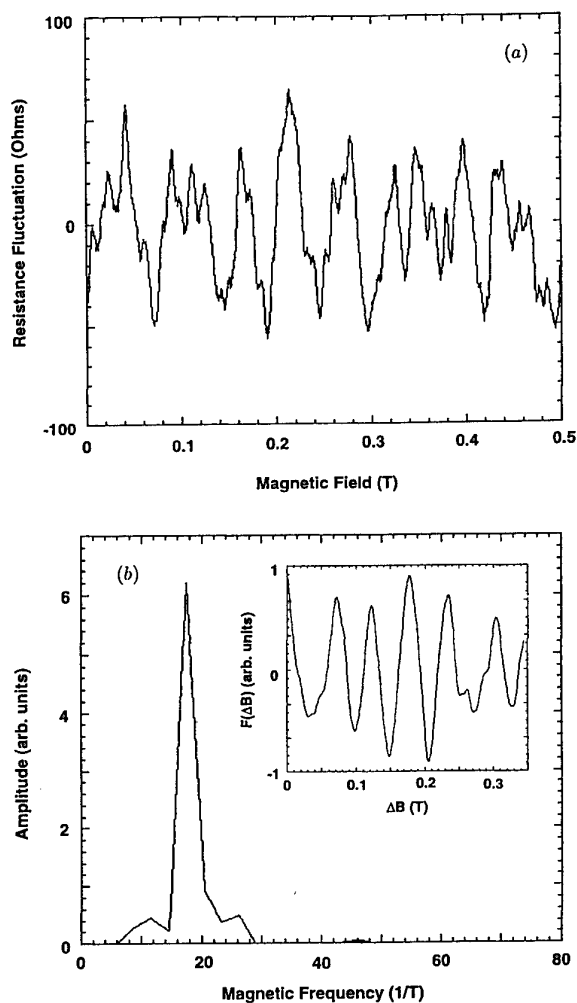


Figure 4. Quantum magnetotransport in a ballistic quantum dot in an InGaAs quantum well system. The fluctuating part of the magnetoresistance is shown in (a) with the actual in-plane gated device above. The correlation function of the fluctuations is plotted as an inset in (b) along with the Fourier transform showing the single strong peak with weaker subsidiary peaks.

surface Schottky gates [19]. We have used such structures to study the role of edge scattering in the quantum wires,

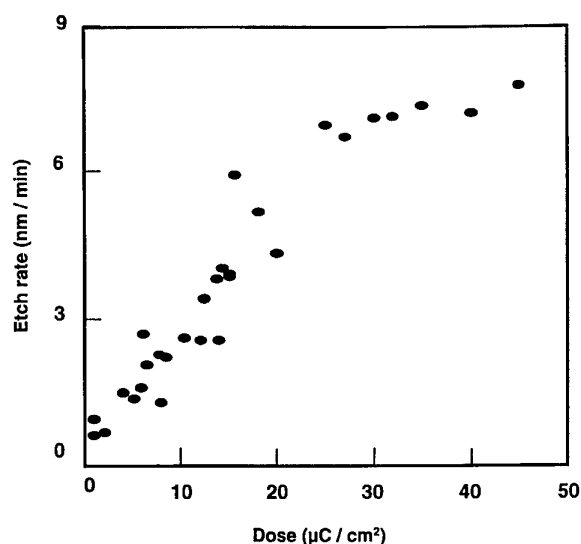


Figure 5. The etch rate of SiO_2 in hot HF vapour for a given electron-beam exposure dose is plotted for a 40 kV exposure.

since when the wire width is smaller than the bulk mean-free path, edge scattering dominates and produces a new mean-free path $\lambda \sim \alpha W$, where W is the wire width. By studying the resistance shown in figure 3, we find that there is a clear transition to edge scattering and we estimate that α is about 3.4 for these wires.

In the top part of figure 4, we illustrate another quantum structure, which is a ballistic quantum dot in which the transport takes place in an InGaAs quantum well. The resistance fluctuations shown in figure 4(a) are thought to arise from the interference of a few regular orbits which are quantized versions of the persisting classical orbits [20].

The correlation function of the fluctuations is shown in the inset to figure 4(b), and the few regular periodic oscillations persist here [21]. This is supported by the existence of only a few sharp frequencies in the Fourier transform of figure 4(b). In chaotic quantum dots, the spectrum is rather broad without these sharp features. Interference is thought to occur at the input quantum point contact [22, 23], with the overall conductance determined by the pair of point contacts.

4. Silicon dioxide as a high-resolution resist

There has been early work on the use of silicon dioxide as a resist by the group at Cambridge [24], but this used electron-beam damage to expose the resist, which required intense dosages. In contrast, through the use of a high-temperature HF vapour etch as a developer, it is possible to use SiO_2 as a high-resolution resist with sensitivity comparable to PMMA [25]. The presence of surface hydrocarbons and the high-temperature HF vapour provide catalytically enhanced etching for the development process. Exposure of the SiO_2 is accomplished either by electron-beam irradiation or by STM exposure. These patterned films can then be used as masks for subsequent processing and pattern transfer.

Exposure with an electron-beam system is accomplished by the fixing of carbon at the exposure site, and this process can produce a positive resist with lines as fine as 20 nm. The use of the STM for exposure has demonstrated even higher resolution, with features as small as 5 nm being developed by RIE through the developed SiO_2 resist. These results are due solely to the presence of the hydrocarbon pattern created by the electron beam and not to either a contamination effect or to oxide damage. In figure 5 the etch rate of the SiO_2 for a given exposure dose

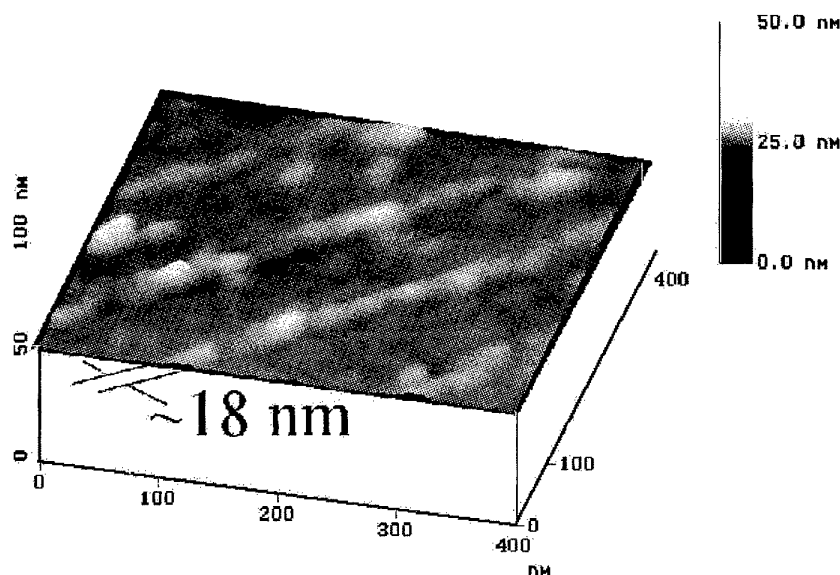


Figure 6. Lines of CoSi_2 formed by STM exposure of SiO_2 , followed by hot HF development, deposition and silicidation of Co, and finally removal of unreacted Co and remaining SiO_2 .

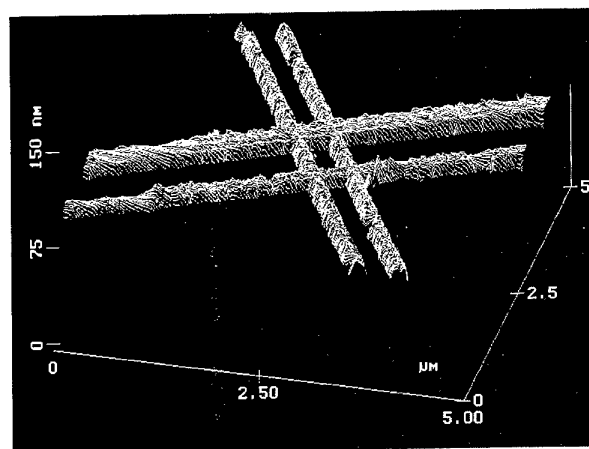


Figure 7. An AFM image of a negative pattern in a Cr film after the background Cr is etched away in a standard etchant leaving the low-dose oxidized regions. The oxide lines are 100 nm wide and are separated by about 100 nm.

is plotted for a 40 kV exposure. The etch rate is sensitive to both dose rate and electron beam energy. A 20 nm Au line has been produced by lift-off. In figure 6, CoSi_2 lines are shown that have been produced by exposure, deposition of a Co film, silicidation by rapid thermal processing, and then removal of the remaining Co and SiO_2 . Continuous lines with a width of 12–15 nm have been fabricated by this method, although individual feature sizes much smaller can be obtained. More recently, the use of monolayer Langmuir–Blodgett films as a source for the C species has been investigated.

We have also used self-assembled monolayers of *n*-octadecyltrichlorosilane, hexadecanol, myristic acid and undecylenic acid on the surface of the SiO_2 thin films as a more robust source of carbon during the exposure process. Infrared analysis of these monolayers supports the formation of a carboxylic acid in the self-assembly process by the hydrogen bonding coordination of the acid hydroxyl group to surface SiOH bonds. Experiments comparing the etching behaviour of the SiO_2 films coated with these materials and directly with carboxylic acids supports this result. We find that etching is not initiated by the backbone carbon of these films but by the interfacial carbon. These films provide an additional layer of control when using the SiO_2 films for the resist in either the e-beam or STM patterning.

5. STM-based lithography

Early in the development of the scanning-tunnelling microscope it was realized that the atomic resolution micrographs that were being obtained suggested that very high resolution lithography could also be achieved. Surprisingly, while the operating voltage is much smaller (5 V versus 40 kV), the beam currents in STM lithography are orders of magnitude higher than the normal electron-beam source, and the STM has been used extensively in this arena. At ASU, the STM has been used to expose

resists and to induce selective area oxidation of both Si and Cr structures [26], as well as Si_3N_4 .

Experiments on the oxidation of Cr are usually begun by depositing Cr on clean Si surfaces in a high-vacuum environment. Cr was selected because of its documented multistates of oxidation, as well as providing a conducting path for the tunnelling electrons. The Cr surface was then modified by the STM patterning. The patterned surface could then be selectively etched to produce either a positive or negative pattern (see figure 7), depending upon the dose delivered to the Cr by the STM tip. That is, it was found that a low dose usually meant that a positive pattern would form, while a high dose would produce a negative pattern. The red colour and water solubility of the patterns formed with a high dose suggests that the Cr has been oxidized to a high oxidation state, possibly CrO_3 . For the low-dose modified regions, the strong chemical and physical properties suggest that a relatively low oxidation state of Cr has been formed, possibly Cr_2O_3 . Auger electron spectroscopy of the modified regions supports these conclusions. In figure 7, a pattern formed by high-dose oxidation of the Cr film is illustrated.

6. Conclusions

Nanolithography can be used to define critical dimensions in ultrasmall structures to below 10 nm by a variety of approaches. However, one cannot simply begin with one aspect in isolation; rather, the proper approach must combine exposure tool, resist, development and pattern transfer techniques into a synergistic methodology. When this is done, one can create truly novel approaches to microdevice fabrication [27].

Acknowledgments

This work supported in part by the Advanced Research Projects Agency and the Office of Naval Research.

References

- [1] Bernstein G and Ferry D K 1986 *Superlatt. Microstruct.* **2** 373
Ryan J, Han J, Krizan A, Ferry D K and Newman P 1989 *Solid State Electron* **32** 1609
Han J, Ferry D K and Newman P *IEEE Electron Device Lett.* **11** 209
- [2] Jin Y, Mailly D, Carcenac F, Etienne B and Launois H 1987 *Microelectron. Eng.* **6** 195
- [3] Ishibashi A, Funato K and Mori Y 1988 *Japan. J. Appl. Phys.* **27** L2382
- [4] Ono M, Saito M, Yoshitomi T, Fiegna C, Ohguro T and Iwai H 1993 *Proc. IEDM* (New York: IEEE Press) p 119
Ono M, Saito M, Yoshitomi T, Fiegna C, Ohguro T, Momse H S and Iwai H 1995 *J. Vac. Sci. Technol. B* **13** 1740
- [5] See, for example, Chamberlain J M, Eaves L and Portal J-C 1990 *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures* (New York: Plenum)
Ferry D K, Barker J R and Jacoboni C 1991 *Granular Nanoelectronics* (New York: Plenum)

- Ferry D K, Grubin H L, Jacoboni C and Jauho A-P 1995 *Quantum Transport in Ultrasmall Devices* (New York: Plenum)
- [6] See, for example, Kelly M J and Weisbuch C (ed) 1986 *The Physics and Fabrication of Microstructures and Microdevices* (Berlin: Springer)
- [7] McCord M A and Pease R F W 1988 *J. Vac. Sci. Technol. B* **4** 86
- [8] Marrian C R K, Dobisz E A and Dagata J A 1992 *J. Vac. Sci. Technol. B* **10** 2877
- [9] Day H C, Allee D R, George R and Burrows V A 1993 *Appl. Phys. Lett.* **62** 1629
- [10] Majumdar A, Oden P I, Carrejo J P, Nagahara L A, Graham J J and Alexander J 1992 *Appl. Phys. Lett.* **61** 2293
- [11] Bernstein G, Ferry D K and Liu W-P 1989 *US Patent* 4877 716; 1990 *US Patent* 4937 174
Bernstein G, Hill D A and Liu W-P 1992 *J. Appl. Phys.* **71** 4066
- [12] Howard R E and Prober D E 1982 *VLSI Electronics: Microstructure Science* vol 5 ed N G Einspruch (New York: Academic) p 146
- [13] Ferry D K, Bernstein G and Liu W-P 1988 *Physics and Technology of Submicron Structures* ed H Heinrich, G Bauer and F Kuchar (Heidelberg: Springer) p 37
- [14] Broers A N 1981 *J. Electrochem. Soc.* **128** 166
- [15] Craighead H G, Howard R E, Jackel L D and Mankiewich P M 1983 *Appl. Phys. Lett.* **42** 38
- [16] Chen W and Ahmed H 1993 *J. Vac. Sci. Technol. B* **11** 2519
- [17] Khoury M and Ferry D K 1996 *J. Vac. Sci. Technol. B* **14** 75
- [18] Ma J, Puechner R A, Liu W-P, Mezenner R, Kriman A M, Maracas G N, Ferry D K, Chu P and Wieder H H 1989 *Proc. InP and Related Compounds for Advanced Electrical and Optical Devices* ed R Singh and L J Messick (Bellingham, WA: SPIE) (*SPIE Proc.* **1144** 619)
- Ma J, Puechner R A, Kriman A M and Ferry D K 1990 *Localization and Confinement of Electrons* ed F Kuchar, H Heinrich and G Bauer (Berlin: Springer) p 268
- [19] Pivin D P Jr and Ferry D K *Physica B* at press
- [20] Bird J P, Ferry D K, Akis R, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Europhys. Lett.* at press
- [21] Connolly K M, Pivin D P Jr, Ferry D K and Wieder H H *Superlatt. Microstruct.* at press
- [22] See, for example, Berry M V and Tabor M 1976 *Proc. R. Soc. A* **349** 101; 1977 *J. Phys. A: Math. Gen.* **10** 371
- [23] Grincwajg A, Edwards G and Ferry D K 1996 *Physica B* at press
- [24] Allee D R and Broers A N 1990 *Appl. Phys. Lett.* **57** 2271
Allee D R, Umbach C P and Broers A N 1991 *J. Vac. Sci. Technol. B* **9** 2839
- [25] Whidden T K, Allgair J, Ryan J M, Kozicki M N and Ferry D K 1995 *J. Electrochem. Soc.* **142** 1199
Whidden T K, Allgair J, Jenkins-Gray A and Kozicki M N 1995 *J. Vac. Sci. Technol. B* **13** 1337
Whidden T K, Allgair J, Jenkins-Gray A, Khoury M, Kozicki M N and Ferry D K 1995 *Japan. J. Appl. Phys.* **34** 4420
- [26] Song H J, Rack M J, Abugharbieh K, Lee S Y, Khan V, Ferry D K and Allee D R 1994 *J. Vac. Sci. Technol. B* **12** 3720
- [27] See, for example, Campbell P and Snow E S 1996 *Semicond. Sci. Technol.* **11** 1558

Proximal probe-based fabrication of nanostructures

P M Campbell and E S Snow

Naval Research Laboratory, Washington, DC 20375, USA

Abstract. We describe a simple and reliable process for the fabrication of nanometre-scale silicon structures by using the local electric field of a proximal probe tip (either an STM or a conducting tip atomic force microscope) to write surface oxide patterns by local anodic oxidation. These oxide patterns can be used as masks for selective etching to transfer the pattern into the substrate. This process has been used to fabricate side-gated Si field effect transistors with critical features as small as 30 nm. Alternatively, this process of anodic oxidation can be used to oxidize completely through thin metal films to make lateral metal–oxide–metal tunnel junctions.

1. Introduction

Originally developed to image surfaces with atomic precision, proximal probes such as the scanning tunnelling microscope (STM) and the atomic force microscope (AFM) have recently been found capable of modifying surfaces at the nanometre scale [1]. In what represents the ultimate limit of surface patterning, an STM has been used to manipulate and position single atoms on a surface [2]. Such a level of control over the surface suggests that proximal probes may be useful in the fabrication of nanostructures. Various approaches have been used to generate patterns useful for device fabrication, including STM tip-induced current exposure of polymeric resists [3], direct writing of metal patterns on a surface by the STM tip-induced decomposition of an organometallic compound [4] or by field evaporation of metal atoms directly from a STM tip [5], among many others. However, these techniques have suffered from numerous problems, including low write speeds and poor reproducibility. A basic difficulty arises because surface patterns generated by proximal probes are by definition thin, and hence are often not sufficiently robust to withstand subsequent processing steps.

In particular, the exposure of polymeric resists (the standard patterning method used in microelectronic device fabrication) by use of an electrically biased proximal probe tip will be limited by the fact that (in contrast to a focused electron beam) a biased tip separated from the substrate surface by a distance equal to any usable thickness of resist (at least a few tens of nanometres, but often considerably more) is a divergent point source of current or electric field. This should be contrasted with an STM tip in true tunnelling mode, which requires a tip–substrate separation of at most a few tenths of a nanometre. In this case the strong exponential dependence of the tunnelling mechanism serves to sharpen the current flow and renders it effectively convergent. Therefore, while focused electron beams can easily expose relatively thick layers, biased proximal

probe tips cannot do so without sacrificing resolution. Unfortunately, thin layers of resist are unsuitable for many standard processing steps, including metal lift-off (which usually requires a sacrificial resist layer several times the thickness of the layer to be patterned) or selective etching (thin resist patterns often cannot withstand any but the most selective of etch techniques) necessary to fabricate usable structures. What is therefore needed is a proximal probe-based surface modification technique which produces a thin (for high resolution) modified layer that can withstand pattern transfer techniques to produce useful nanostructures.

2. Proximal probe-induced local oxidation

A proximal probe-based surface modification technique that shows great promise is voltage-induced surface oxidation. Early work on this was done by Dagata [6], who used an STM tip operated in air to oxidize local regions of a hydrogen-passivated Si (111) surface. This surface oxidation process is attractive because the local oxide pattern so written, which can be as thin as one monolayer, is distinct from the substrate from the standpoint of both chemical and electronic properties.

We have devised a method to utilize such surface oxide patterns to fabricate operating nanometre-scale devices. This approach uses the oxide pattern as an etch mask for selective etching of the pattern into the underlying substrate. This process requires a selective etch which attacks the substrate material not protected by the proximal probe-generated surface oxide pattern but does not etch the oxide pattern, or at least attacks it very slowly in comparison with the underlying material. The oxidation process is similar to that of Dagata [6] in that a biased proximal probe tip operating in air locally oxidizes a hydrogen-passivated silicon surface. However, we use the Si (100) surface instead of the Si (111) used by Dagata because the Si (100) surface etches much faster in certain liquid etches than does the Si (111), thus increasing the etch selectivity.

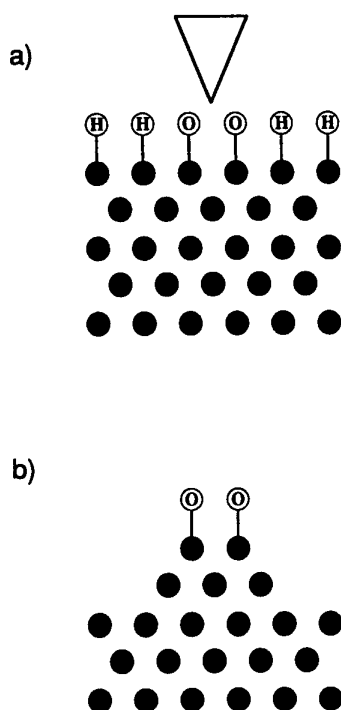


Figure 1. Schematic diagram showing (a) the proximal probe tip locally oxidizing a small region of an H-passivated Si surface and (b) use of this oxide pattern as a mask for selective liquid etching. Liquid etches generally have etch rates that vary with crystal direction, and thus expose crystalline planes with low etch rates. As these low-etch-rate planes are exposed, there will be some undercutting of the oxide mask (not shown in the diagram).

Our process is simple and reliable. The surface of the Si (100) is first immersed in a weak ($\sim 10\%$) aqueous solution of HF for approximately 1 min and then blown dry. This immersion strips the native oxide layer from the surface and replaces it with a monolayer of hydrogen. This hydrogen monolayer terminates all the dangling surface bonds and renders the surface chemically stable. The surface will remain passivated over a period of days or longer, although exposure to heat or certain chemicals such as hydrocarbons can destroy the passivation and expose the surface to reoxidation in air. The passivated silicon is then ready for selective oxidation with a scanning proximal probe (either an STM [7] or a conducting tip AFM [8]). As shown in figure 1(a), the application of a bias of a few volts between tip and passivated silicon surface creates a high local electric field which in air induces anodic oxidation of the surface in the immediate vicinity of the tip. The oxide thickness is primarily a linear function of tip bias, producing a few angstroms of oxide per volt of tip bias. The oxide thickness can therefore be easily and reliably controlled over the range from one monolayer up to several nanometres. Lower humidity generally gives narrower lines, but at humidity much below $\sim 20\%$ it is difficult to write an oxide pattern. This humidity dependence demonstrates that water vapour is necessary for the oxidation process. In air the smallest oxide patterns we have been able to write are ~ 10 nm diameter dots and

~ 10 nm wide oxide lines. This resolution limit of ~ 10 nm in air appears to be set not by intrinsic tip sharpness but by the need for a certain minimum level of ambient humidity.

While this tip-induced oxidation can be performed with an STM [7], there are several advantages to using instead a conducting tip AFM [8]. With STM, the tip bias is required both for the oxidation mechanism and for feedback control of the imaging. With AFM, however, the imaging is done by applied force as measured by the deflection of the cantilever supporting the tip. The imaging mechanism for AFM (contact force) is thus independent of the oxidation mechanism (tip bias). This implies that with a conducting tip AFM one can image an oxidized pattern with the same tip used to do the oxidation without fear of further oxidation. With an STM, imaging of an oxidized pattern must be done carefully in order not to oxidize the surface further. Many commercially available AFM tips are not electrically conducting. We have found that a 30 nm thick layer of titanium deposited on the AFM tip by electron beam evaporation will render such tips electrically conducting. The Ti film is mechanically strong enough to withstand repeated scanning for imaging without abrading or tearing, and can also tolerate the voltages necessary for oxidation without melting or vaporizing. These failure modes are common with other metal coatings such as Au and Pt. We attribute the robust nature of the Ti tips to a thin surface oxide that forms in the surface, which allows the application of electrical bias but prevents significant current flow from the tip. We detect no measurable current flow during the oxidation process. The oxidation mechanism is therefore mediated by the strong local electric fields generated by the biased tip and not by a large electronic current flow from the tip.

Both the H-passivated Si surfaces and the oxide patterns written on them as described above are quite stable and can last without pattern degradation for several days in air with normal handling (provided they are not exposed to heat or chemical contamination such as hydrocarbons). The oxide patterns can be used as an etch mask for pattern transfer into the silicon, as illustrated in figure 1(b). There exist several liquid etches that etch silicon but not its oxides, or etch its oxides very slowly. Two such etches are hydrazine and aqueous solutions of KOH. Hydrazine is completely selective against etching even a single monolayer of oxide and (although highly toxic and flammable) is hence useful for deep etching using thin oxide masks. KOH solutions do attack oxides of silicon, but at rates several orders of magnitude below the rate at which it attacks silicon [9]. This selectivity of KOH is sufficient for all but the deepest of etches with the thinnest of oxide masks. These and similar liquid etches do not etch vertical sidewalls but expose those crystalline planes with low etch rates, as shown in figure 1(b). In order to get vertical sidewalls, directionally selective gas etches must usually be employed.

3. Device fabrication

In order to fabricate useful devices (for example, narrow conducting wire channels and the more complex structures such as transistors that can be made from them) the

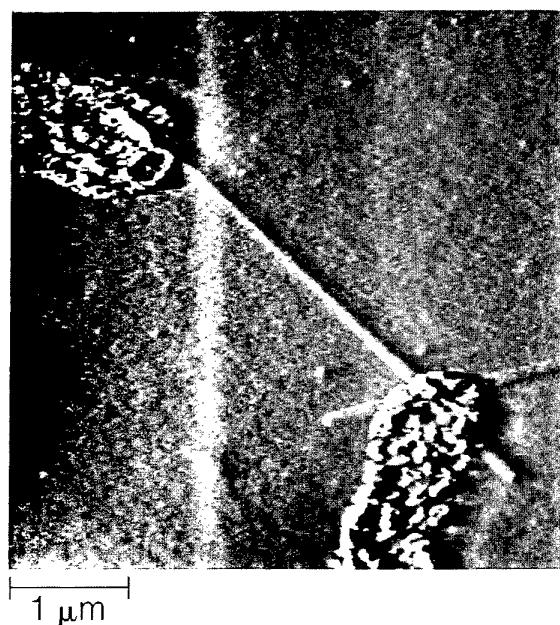


Figure 2. AFM gray-scale image of an oxide line written between two metal contact pads. The imaging was done immediately after the pattern was written by the same tip used to write the pattern. Although the oxide pattern is only 2 nm high by 50 nm wide, it is clearly visible by AFM. Any breaks or defects can be repaired before pattern transfer by selective etching.

completed devices must be electrically isolated from the substrate. The oxide patterning and etching described above must therefore be done on a thin layer of Si sitting on an insulating layer. For this purpose we use SIMOX, which is formed by ion implanting a bulk silicon wafer with a heavy dose of oxygen and annealing it to form a buried layer of SiO₂ sandwiched between a thin surface Si layer and the bulk of the Si. The formation and electrical properties of SIMOX are discussed in detail elsewhere [10].

The first step is to pattern on the surface of the SIMOX wafer by standard optical lithography and lift off large metal bond pads and ohmic contacts with fingers extending down to the relatively small ($\sim 10 \mu\text{m} \times \sim 10 \mu\text{m}$) active region to be patterned by proximal probe. The surface of the sample is next passivated with dilute HF as described above. The sample is then mounted in the proximal probe for patterning. Because of the many advantages that an AFM offers over a STM for this type of patterning (in particular the ability to image without exposure), we prefer for this step a conducting tip AFM, as described above.

An AFM image is next taken of the active area between the contact fingers. This image is used to register the tip to the active area to be patterned. To write the oxide pattern for (as an example) an isolated conducting wire between two contacts, the tip is first moved to one of the pads, the oxidizing bias (a few volts, usually in the range of 2 to 8 V, tip negative with respect to the sample) is turned on, and the tip is scanned (typical scan speed is $1\text{--}10 \mu\text{m s}^{-1}$, although write speeds of 1 mm s^{-1} or higher should be possible [8]) from one pad to the other to define the wire pattern. Upon reaching the other contact pad, the oxidizing bias is then

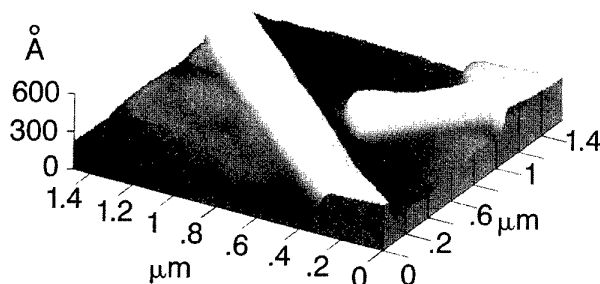


Figure 3. AFM image of the source-drain channel and side gate of an Si FET patterned by AFM and etched in hydrazine.

turned off. At this point the oxide pattern of the wire can be imaged for accuracy and continuity. Figure 2 shows an AFM image of an oxide line written between two pads. The imaging was done immediately after the pattern was written by the same tip used to write the pattern. The metal contact fingers are 20 nm high by $2 \mu\text{m}$ wide, while the oxide pattern, although only 2 nm high by 50 nm wide, is clearly visible. Any defects in the pattern can be found and repaired before the pattern is transferred by etching to form the silicon wire. A common defect is incomplete oxidation in the region where the wire joins the contact finger, which will arise if the tip rides over the edge of the contact finger too abruptly and does not provide sufficient exposure at the joining point. It is a trivial matter to find these defects and repair by additional oxidation the appropriate areas before the pattern is etched.

The etching is done with warm (70°C) hydrazine. Hydrazine is preferable in this case to KOH because K contamination may affect the electrical properties of the Si-SiO₂ interface. The hydrazine etches all the H-passivated Si surface but not that masked by the AFM-grown oxide pattern or protected by the contact and finger metal patterns. The underlying SIMOX oxide acts as an etch stop. The net result in etching the pattern shown in figure 2 will be an Si wire connected between two contact pads and resting on an insulating layer of SiO₂, which itself rests on an underlying Si layer. Wires fabricated by this technique are electrically conducting and, because they rest on a layer of SiO₂, an appropriate voltage applied to the underlying Si layer can drive the wire into accumulation, depletion or inversion, thus giving independent control over the conductance and carrier type of the wire [11].

The conducting wire between two pads is the simplest transport structure that can be made and is the basic building block for more complex devices. A side-gated field effect transistor (FET) can be made from the wire pattern in figure 2 by adding a side gate consisting of a wire that connects to a third contact pad and approaches (but does not touch) the source-drain channel (the wire between the two pads). Figure 3 shows an AFM image of a completed side-gated FET structure fabricated by this process. The source-drain channel wire and side gate, both resting on the SIMOX SiO₂ layer, are clearly visible. The source-drain channel wire is $\sim 200 \text{ nm}$ wide by 40 nm high. The side gate has similar dimensions but is separated from the source-gate channel by 60 nm.

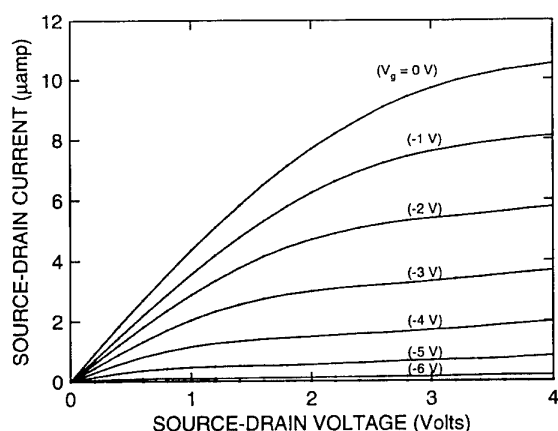


Figure 4. Source-drain current versus source-drain voltage for various gate biases for a side-gated Si FET fabricated by AFM.

Because the source-drain channel of such devices has such a small cross-sectional area, the channel conductance is quite small, even for relatively heavily doped material (typical doping is 10^{15} – 10^{17} cm^{-3} , n-type). The application of an accumulating bias to the substrate as described above can significantly raise the conductance of the source-drain channel. The side gate can then be used to modulate the channel conductance by depleting the channel and restricting the current flow from source to drain. Figure 4 shows the transistor characteristic of a device similar in structure to that shown in figure 3 (channel width is 300 nm, channel height is 200 nm, doping is 10^{17} cm^{-3} n-type) backgated +40 V into accumulation and depleted by a single side gate [7]. The side gate bias ranges from 0 V to –6 V in –1 V steps. Operating transistors with source-drain wires as thin as 40 nm and narrow as 30 nm with gate-channel separations as small as 25 nm have been fabricated by this technique. Such small devices can, by backgating into accumulation, achieve currents as high as 5×10^{-7} A [12].

These techniques can be used to fabricate other types of devices. Recently, an otherwise conventional large-scale Si MOSFET was fabricated, in which a $0.1 \mu\text{m}$ gate made of amorphous hydrogenated Si was defined by AFM-induced oxidation and dry etching [13]. Dry rather than liquid etching was used to etch the amorphous silicon gate because the lack of crystalline structure precludes the well-defined etch profiles achieved with liquid etches on single-crystal material. Other materials such as GaAs can be patterned by depositing a thin film of Si which, when patterned and etched, acts as a resist to pattern the underlying substrate [14].

4. Recent developments and future directions

The techniques described above use the anodic oxide pattern written by the proximal probe as an etch mask to transfer the pattern into the substrate. An alternative approach is to use the oxide pattern directly as an active element of the device structure. Sugimura *et al*

[15] demonstrated tip-induced anodic oxidation of thin Ti films. Matsumoto *et al* used a conducting tip AFM to oxidize completely through thin Ti films and thus produce lateral metal-oxide-metal junctions [16]. More recently, Matsumoto *et al* used these Ti–TiO_x lateral junctions to fabricate a single-electron tunnelling device that shows Coulomb blockade effects at room temperature [17]. The smallest TiO_x patterns were of the order of 20–30 nm wide, which is roughly comparable with the minimum size achieved by oxidizing and etching Si as described above. In order to reach sizes below this scale, a more controlled approach will be needed.

We have added additional control to the above process by monitoring the electrical resistance of the Ti film *in situ* during the selective oxidation [18]. By measuring the resistance of Ti wire (starting width $\sim 2 \mu\text{m}$ wide by 7 nm thick) as its width was being effectively reduced by conducting tip AFM oxidation and by terminating the oxidation when the appropriate resistance (and hence the desired width) was reached, we were able to produce conducting Ti wires as narrow as ~ 5 nm. This same technique of *in situ* measurement and feedback control of resistance allows us to form Ti–TiO_x–Ti junctions of precisely controlled resistance. This process of *in situ* measurement and control of electrical properties should allow the fabrication of devices with critical dimensions below 10 nm, a region in which operating characteristics will depend critically upon geometry but for which control of these characteristics by controlling geometry alone would be extremely difficult.

To achieve resolution down to 1 nm and below, one will most likely have to go into a more controlled environment than ambient. Lyding *et al* [19] recently used an STM operating in UHV to depassivate selectively a H-passivated Si surface and so write 1 nm wide lines of depassivated surface on a 3 nm pitch. Exposure of similar patterns to atomic hydrogen has produced oxide lines on an 8 nm pitch [19]. These results, along with the atomic-level resolution already achieved by proximal probe imaging and single-atom manipulation, suggest that proximal probe lithography may soon achieve the ultimate limit of atom-by-atom control of device fabrication. While the above results have all been achieved with a single tip writing serially, recent progress in the micromachining of multiple AFM tips on a single chip offer the possibility of eventually writing multiple patterns in parallel [20]. The integration of a large number of independent AFM tips on a single chip may eventually make massively parallel AFM lithography an attractive candidate for large-scale commercial nanofabrication in the size regimes well below the limits of optical and perhaps electron beam lithography.

References

- [1] Wiesendanger R 1992 *Appl. Surf. Sci.* **54** 271
- [2] Eigler D M and Schweizer E K 1990 *Nature* **344** 524
- [3] McCord M A and Pease R F W 1988 *J. Vac. Sci. Technol.* **B 6** 293
- [4] Silver R M, Ehrichs E E and de Lozanne A L 1987 *Appl. Phys. Lett.* **51** 247
- [5] Lyo I-W and Avouris P 1991 *Science* **253** 173

- [6] Dagata J A, Schneir J, Harary H H, Evans C J, Postek M T and Benett J 1990 *Appl. Phys. Lett.* **56** 2001
- [7] Snow E S, Campbell P M and McMarr P J 1993 *Appl. Phys. Lett.* **63** 749
- [8] Snow E S and Campbell P M 1994 *Appl. Phys. Lett.* **64** 1932
- [9] Palik E D, Glembocki O J and Stahlbush R E 1988 *J. Electrochem. Soc.* **135** 3126
- [10] McMarr P J, Mrstik B J, Barger M S, Bowden G and Blanco G R 1990 *J. Appl. Phys.* **67** 7211
- [11] Campbell P M, Snow E S and McMarr P J 1994 *Solid-State Electron.* **37** 583
- [12] Campbell P M, Snow E S and McMarr P J 1995 *Appl. Phys. Lett.* **66** 1388
- [13] Minne S C, Soh H T, Flueckiger Ph and Quate C F 1995 *Appl. Phys. Lett.* **66** 703
- [14] Snow E S, Campbell P M and Shanabrook B V 1994 *Appl. Phys. Lett.* **63** 3488
- [15] Sugimura H, Uchida T, Kitamura N and Masuhara H 1993 *Japan. J. Appl. Phys.* **32** L553
- [16] Matsumoto K, Takahashi S, Ishii M, Hoshi M, Kurokawa A, Ichimura S and Ando A 1995 *Japan. J. Appl. Phys.* **34** 1387
- [17] Matsumoto K, Ishii M, Segawa K, Oka Y, Vartanian B J and Harris J S 1996 *Appl. Phys. Lett.* **68** 34
- [18] Snow E S and Campbell P M 1995 *Science* **270** 1639
- [19] Lyding J W, Shen T-C, Hubacek J S, Tucker J R and Abein G C 1994 *Appl. Phys. Lett.* **64** 2010
- [20] Minne S C *et al* 1995 *Appl. Phys. Lett.* **67** 3918; 1995 *J. Vac. Sci. Technol. B* **13** 1380

STM investigation and manipulation of C₆₀ molecules adsorbed on an Si(111) surface

A W Dunn, P Moriarty, M D Upward, A Nogaret and P H Beton

Department of Physics, University of Nottingham, Nottingham NG7 2RD, UK

Abstract. We have used an ultra-high-vacuum (UHV) scanning tunnelling microscope (STM) to image C₆₀ molecules adsorbed on a Si(111) 7 × 7 surface. At low coverage (~0.01 monolayers) molecules are adsorbed at random sites. For coverages close to a monolayer they are partially ordered in a hexagonal arrangement. Second- and higher-layer islands, in which the C₆₀ molecules are clearly resolved, are observed at higher coverage. These islands may be desorbed by annealing in the range 200–300 °C, leaving an Si surface terminated by a C₆₀ monolayer. This surface is stable to exposure to air and immersion in water. In addition, recent work on manipulation of C₆₀ molecules at various coverages is reviewed and results relating to tip alignment in UHV are discussed.

The use of scanning probe microscopes as tools for patterning surfaces offers a means of fabricating experimental device structures which have an active region comparable with single atoms or molecules. There are several different ways of using the tip of a scanning tunnelling microscope (STM) to pattern a surface. In the first demonstration of atomic-scale modification by Becker *et al* [1] an atom was transferred from tip to substrate by applying a voltage pulse. The most celebrated example of atomic manipulation was achieved by exploiting the attractive force between an STM tip and an Xe atom to drag an atom across a metal surface [2]. In this and subsequent work [3] low-temperature operation and a metal substrate were used. Another notable example of STM modification is H desorption from the hydrogen-passivated Si surfaces. This desorption results in local chemical changes such as a change in surface reconstruction [4, 5] for UHV operation or oxide formation for ambient operation [6, 7].

We have recently conducted a series of experiments with the objective of extending the atomic and molecular manipulation techniques described in [2] and [3] to room-temperature operation and the use of semiconductor substrates. At the outset of this work it was not clear that there would be a suitable combination of adsorbate and substrate which would have a sufficiently strong interaction that the arrangement of molecules would remain stable at room temperature, but would be sufficiently weak for the molecules to be moved by the STM tip. C₆₀ on Si(111) 7 × 7 was identified as a promising combination by Li *et al* [8] on the basis of STM results, and also by a number of theoretical groups [9, 10]. In fact, we have found that C₆₀ adsorbed on Si(111) 7 × 7 fulfils these requirements. However, we believe that in our

experiments the interaction between tip and molecule is repulsive, in contrast to the attractive forces which control low-temperature STM manipulation of inert gas atoms [2]. We also draw attention to more recent work by Jung *et al* on manipulation of porphyrin derivatives on a metallic substrate [11].

The Si samples for our experiments were cut from a p-type Si(111) wafer (resistivity > 1 Ω cm). After loading into a UHV system (base pressure 10⁻¹⁰ Torr), the samples were outgassed overnight and flash annealed at 1200 °C for 1–2 min to form a 7 × 7 reconstruction on the sample surface. C₆₀ is deposited from a Knudsen cell at a rate of ~2 monolayers per hour. The Si is held at room temperature for the deposition, and also for STM [12] imaging and manipulation. We use W tips which are electrochemically etched from polycrystalline wire and heated in UHV by electron bombardment.

C₆₀ manipulation may be achieved simply by scanning at an appropriate voltage, *V*, applied to the sample and tunnel current, *I*, while maintaining feedback control of the current, i.e. in the normal imaging mode. This is illustrated in figure 1, which shows two 40 nm × 40 nm images of C₆₀ adsorbed on Si(111) 7 × 7 taken with *V* = -2.5 V and *I* = -0.1 nA. The C₆₀ molecules are the bright topographic features in the image. The Si atoms of the 7 × 7 reconstruction are also clearly resolved. Between the acquisition of these images a 20 nm × 20 nm STM image was acquired (scan direction parallel to the horizontal axis of the image) with *V* = -1.5 V and *I* = 0.1 nA. A reduction in the potential difference between tip and sample will result in a reduction in the tip-sample separation. The area covered by this intermediate scan is identified by the white square in figure 1. It is clear that the C₆₀ molecules

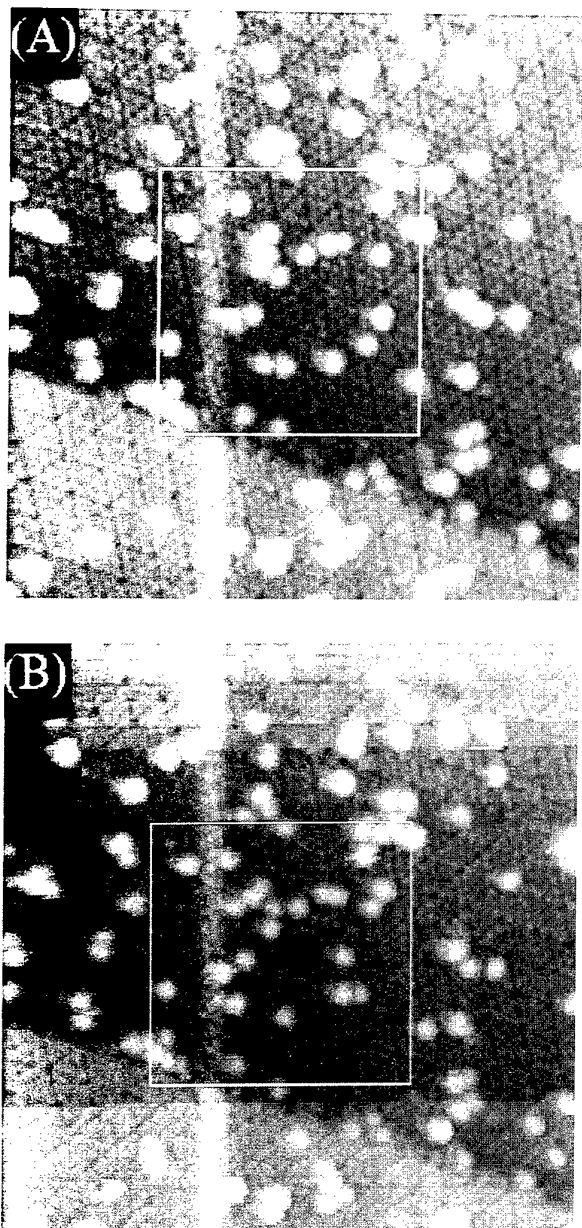


Figure 1. (a) 40 nm² scan showing initial arrangement of molecules prior to reduced bias (sample voltage -1.5 V, 0.1 nA, 20 nm² scan in centre of image—denoted by the white square. The scan parameters (sample voltage, current) were (a) -2.5 V and 0.1 nA. (b) 40 nm², -2.5 V and 0.1 nA scan showing reordering of molecules in central region as a result of a single 20 nm² scan at -1.5 V and 0.1 nA.

outside this white square are unaffected by the intermediate scan but many of those within the square have been moved. This type of manipulation, which we have found difficult to reproduce, results in re-randomization of the arrangement of C₆₀ molecules and was first reported by Maruno *et al* [13]. We have not observed a dependence of the rearrangement on scan direction.

In order to manipulate molecules in a more controlled manner we have used the cycle of tip movements shown

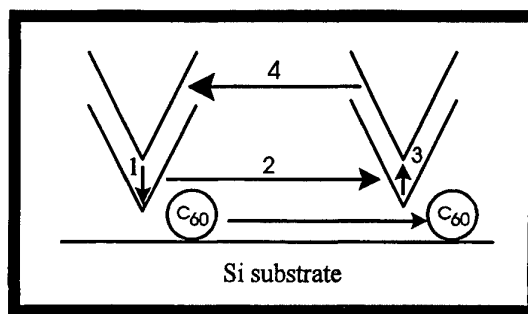


Figure 2. A schematic diagram showing the manipulation procedure used in positioning individual C₆₀ molecules on the Si(111) 7×7 surface.

schematically in figure 2. The tip is first positioned close to a C₆₀ molecule which is to be moved. The tip-sample separation is reduced by adjusting I and V while maintaining feedback control. The tip is moved parallel to the surface and then retracted by readjusting the values of I and V . It is then returned to its initial position. For our experimental arrangement it is difficult to position the tip directly behind the molecule and therefore we typically sweep out a number of parallel lines separated by 0.6 nm. This procedure results in movement of the C₆₀ molecule with a success rate which varies from 10–50%. An example of such movement is shown in figures 3(a)–(c). The tip is first positioned at the cross in figure 3(a) and the sweeping procedure is applied. This results in movement of the uppermost C₆₀ molecule as is clear from a comparison with figure 3(b). The sweeping procedure is then applied again, resulting in further movement of the molecule (see figure 3(c)). Each of the displacements between figures 3(a), (b) and (c) occurred after a single application of the manipulation procedure, although there were also several (typically two to five) unsuccessful applications of the procedure. Typically we find that the molecule does not move through the full extent of the tip excursion during the manipulation procedure. For the movements between figure 3(a)–(c), the tip is moved by 6 nm but the resulting displacements of the C₆₀ are 2 – 3 nm. The C₆₀ is often displaced at an angle (typically $\sim 30^\circ$) to the tip displacement. This, together with the observation that there are few instances of C₆₀ adhesion to the tip, is suggestive of a repulsive interaction between tip and molecule. We have also attempted to move molecules with a positive sample bias, but have found that this produces a significant degree of damage in the substrate surface due to the removal of Si adatoms.

We have used sequences of this procedure to move molecules along two orthogonal directions to create simple patterns [14]. However, this is rather time consuming and we have investigated using this sweeping technique to create patterns in a different way. For this experiment we use a higher coverage of C₆₀ (~ 0.1 monolayer) and rather than attempt to position individual molecules at special sites we try to remove C₆₀ from a particular area. A typical result is shown in figures 4(a)–(c). This is achieved by

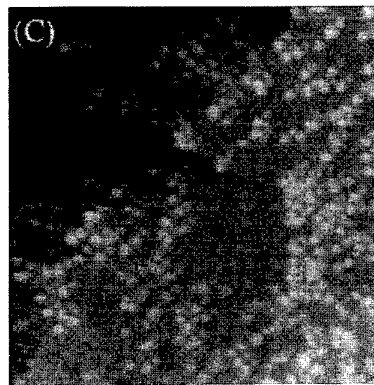
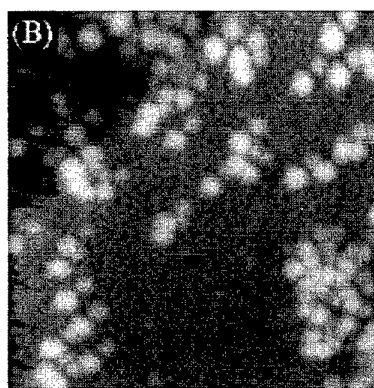
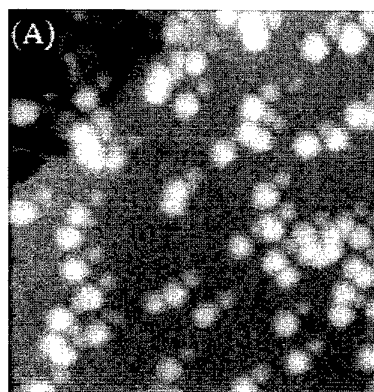
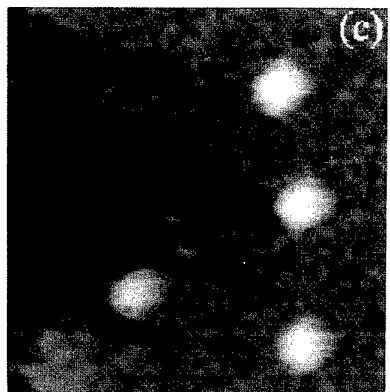
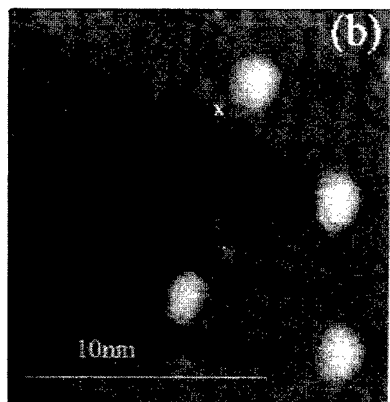
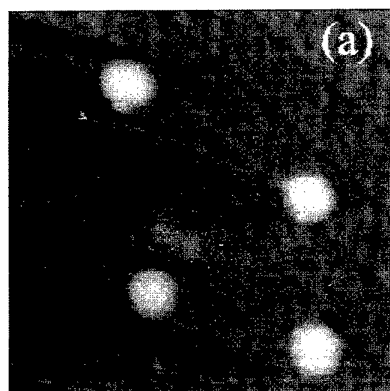


Figure 3. Controlled movement of a single C_{60} molecule on the Si(111) (7×7) surface. The figures are STM topographic images with sample voltage -2.0 V and tunnel current -1 nA. The sweeping parameters were -0.4 V and -4 nA. The molecule in the upper right section of (a) is induced to move approximately 4 nm in the positive x direction, as shown in (b), and subsequently a further 3 nm, resulting in the arrangement of molecules shown in (c). The letter x marks the bottom left corner of the $6 \text{ nm} \times 6 \text{ nm}$ area swept out by the tip in repositioning the C_{60} molecule.

Figure 4. (a) 30 nm^2 scan showing arrangement of molecules after six applications of the sweeping procedure in the lower left of the image. (b) 30 nm^2 scan showing arrangement of molecules after a further 14 applications of the sweeping procedure. (c) 60 nm^2 scan showing a 16 nm^2 region cleared of C_{60} molecules after a total of 31 applications of the sweeping procedure. Sweeping parameters (sample voltage, current): -0.4 V and 4 nA. Scan parameters (sample voltage, current): $+2.0$ V, 0.1 nA.

repeated application of our manipulation procedure. A comparison of figure 4(a) and (b) shows that the molecules in the upper half of figure 4(a) remain unchanged when the manipulation procedure is applied in the lower half of the image. In contrast, the molecules in the bottom half of the image have been moved towards the right-hand side where they form a cluster. Figure 4(c) shows a lower-magnification image in which the area cleared of C_{60} is shown clearly. The local density per unit area of C_{60} molecules is much greater at the right-hand side of the cleared region. The 7×7 reconstruction of the Si surface is resolved in figures 4(a)–(c), and damage to the surface during manipulation is minimal. We have recently repeated this work using a higher coverage of C_{60} (~ 0.25 monolayers). For this coverage, a continuous linear cluster of C_{60} with a width of order 2–3 nm and length 40 nm is formed [15]. An attraction of this approach as compared with positioning individual molecules is that it takes a few minutes rather than a few hours. It is conceivable that by using this technique nanostructures could be fabricated using molecular manipulation which could have dimensions ranging from 1 nm up to several 100 nm, which is comparable with features formed using conventional lithographic techniques.

Figures 5(a) and (b) show STM images after a total coverage of ~ 2 monolayers of C_{60} has been deposited. We find that a complete partially ordered monolayer of C_{60} is formed and then islands of second and third layers are formed with a high degree of hexagonal order. In figure 5(a) the terrace steps of the underlying Si may be resolved. The darkest contrast level is the monolayer coverage, the small bright areas are third-layer islands and the intermediate contrast level regions are second-layer islands. A higher-magnification image shown in figure 5(b) reveals that two domains of hexagonal ordering with principal axes $\pm 11^\circ$ with respect to those of the Si. Such double domain ordering was proposed for thick C_{60} layers by Xu *et al* [16] who proposed that the first monolayer was absorbed at random sites and was therefore disordered. They further proposed that an order–disorder transition took place during the growth of subsequent layers. Chen *et al* [17] showed that there was some short-range order, even in the first layer, but our experiments show that the order can extend over domains of width several tens of nanometres. Note that the nearest-neighbour C_{60} separation in the double domain structure corresponds closely to that observed in bulk C_{60} and so higher layers of C_{60} may be incorporated in well ordered hexagonal layers.

After annealing a sample with the surface structure shown in figure 5(b) in the range 200–300 °C for 1–30 min, we find that the second- and third-layer islands are desorbed, leaving a monolayer termination which retains the double domain ordering. This presumably means that second and subsequent C_{60} layers are bonded rather weakly to the first monolayer. However, the first monolayer is bonded much more strongly to the underlying Si. The C_{60} monolayer therefore provides a surface coating which interacts rather weakly with subsequent adsorbate layers.

To investigate the stability of the C_{60} monolayer-terminated Si(111) 7×7 we have exposed samples to

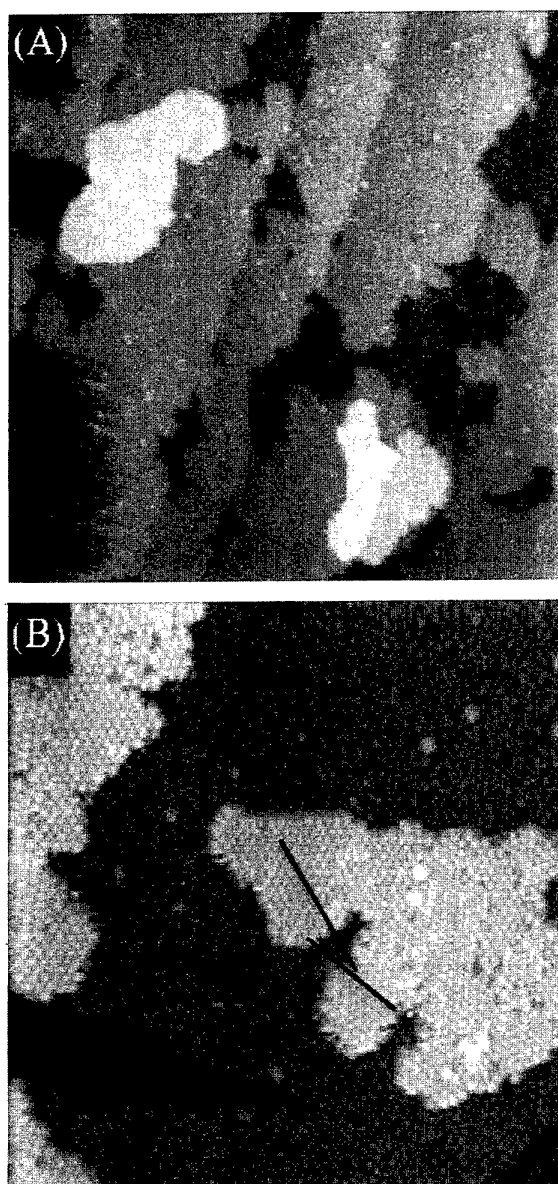


Figure 5. (a) 150 nm² scan of multilayer coverage of C_{60} on the Si(111) 7×7 surface. Dark regions correspond to the complete first monolayer, mid-tone, to patches of second layer, and the brightest contrast regions are third-layer islands. Detail of Si steps can be seen running from the upper right of the image to the lower left region. (b) 60 nm² scan of the surface shown in (a) showing a second-layer island on top of the first monolayer. The thick dark lines show a principal axis for each domain. Scan parameters (sample voltage, current): +3.5 V and 0.1 nA.

atmosphere for up to 30 min and immersed them in water for up to 1 min. This treatment would result in complete corruption of the Si(111) 7×7 surface. In contrast, the C_{60} monolayer appears to completely inhibit chemical attack of Si(111) 7×7 and thus acts to passivate the Si surface. Note that thick C_{60} layers have previously been shown to passivate Si [18] and Al [19] surfaces. Typical results are shown in figure 6. Figures 6(a) and 6(b) show large- and small-area STM scans of a sample which has been exposed

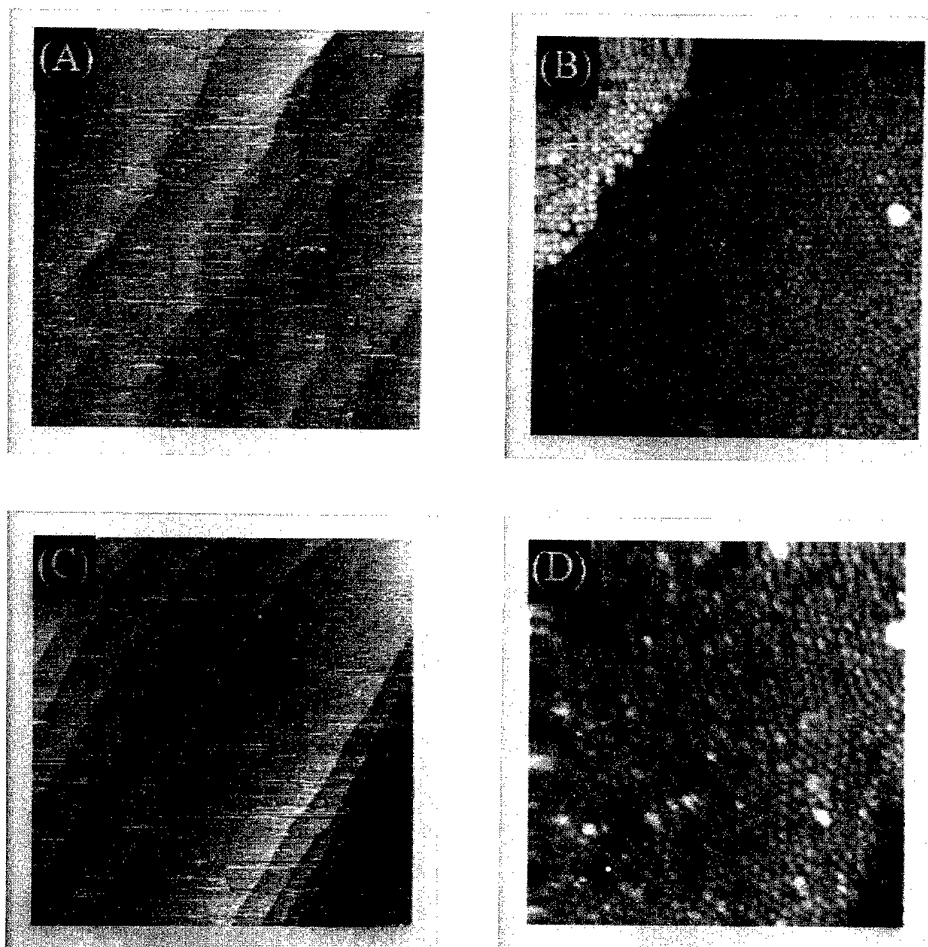


Figure 6. (a) 500 nm² scan of a single monolayer adsorbed on the Si(111) 7 × 7 surface which has been exposed to atmosphere for 1 min. Scan parameters (sample voltage, current): +3.5 V and 0.1 nA. (b) 50 nm² scan of a single monolayer adsorbed on the Si(111) 7 × 7 surface which has been exposed to atmosphere for 30 min, and then annealed at 260 °C for 60 min. Double domain ordering of the C₆₀ molecules is observed. Scan parameters (sample voltage, current): −3.5 V and 0.1 nA. (c) 400 nm² scan of a single monolayer on the Si(111) surface following a 30 s dip in deionized water. Scan parameters (sample voltage, current): +3.5 V and 0.1 nA. (d) 30 nm² scan of the surface shown in (c) following a 260 °C anneal for 15 min. Again, double domain ordering of the first monolayer can be seen indicating that the (7 × 7) reconstruction of the Si is maintained. Scan parameters (sample voltage, current): +3.0 V and 0.2 nA.

to atmosphere for 1 min. This sample has been annealed in UHV at ~ 200 °C for 15 min prior to imaging.

In the large-area image Si terraces are clearly resolved, while in figure 6(b) individual C₆₀ molecules are resolved which are ordered in the double domain structure discussed above. Note that the double domain ordering of the C₆₀ molecules is directly related to the position of the Si adatoms of the underlying 7 × 7 reconstruction. Observation of double domain ordering therefore shows that the underlying 7 × 7 reconstruction of the Si surface remains intact. Note that STM images acquired without annealing also show this hexagonal ordering but image quality is rather poor due, we believe, to physisorbed material on the surface. Similarly, figures 6(c) and (d) show that the C₆₀ molecules and Si terraces can be resolved even after immersion in deionized water for 1 min at room temperature. Thus the C₆₀ monolayer inhibits chemical attack by both water and atmospheric oxygen.

Finally, we address a point which is sometimes neglected in discussions on UHV/STM modification. This relates to the problem of locating specific sites on a surface. This is generally rather difficult since access for optical microscopy is usually limited. One solution is to incorporate the STM in a scanning electron microscope (SEM) [20], although this also has drawbacks in terms of ease of construction and possible corruption of surfaces. As an alternative, we have developed an alignment strategy based on the use of registration marks formed using electron beam and optical lithography. These registration marks may be etched or formed by lift-off of metal overlayers. Figure 7(a) shows an SEM image of some etched marks while figure 7(b) shows STM images of the same area. Note that this central 4 μm × 4 μm square has been located on a 5 mm × 5 mm Si sample. The Si in this case is a (100) wafer, offcut by 4° towards (001). It is interesting to note that after flash annealing we observe Si

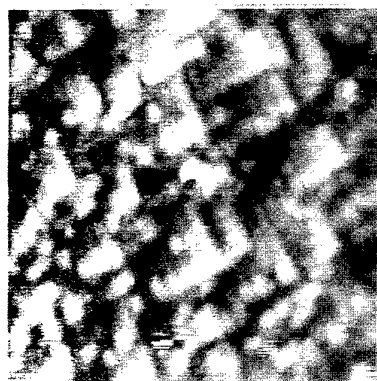
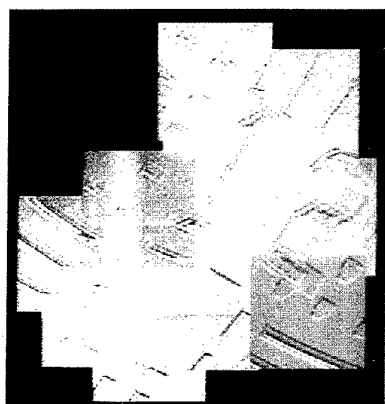
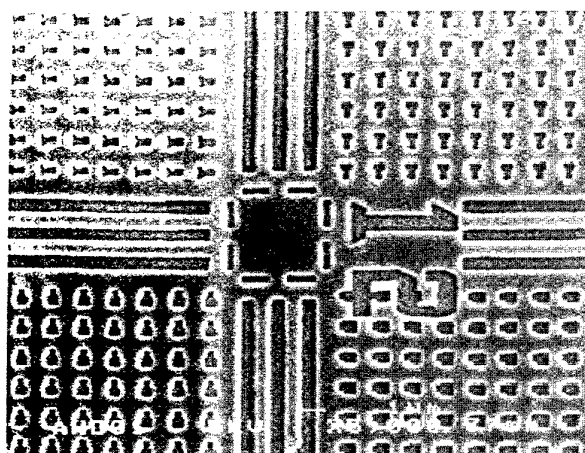


Figure 7. A series of images to demonstrate the formation and use of alignment marks. The top figure shows an SEM micrograph of a series of alignment marks wet etched in a Si(100) (misaligned by 4° towards the (001) direction) surface. These marks were formed using electron beam lithography. The central region is a square of side $4\ \mu\text{m}$. The middle image is a montage of several STM images, each of which covers an area $2.5 \times 2.5\ \mu\text{m}^2$. In addition each of the images has been differentiated to highlight the etched marks. Scan parameters are sample voltage $-4\ \text{V}$, $0.1\ \text{nA}$. The lowest image is taken using the STM (sample voltage $2\ \text{V}$, $0.23\ \text{nA}$, $20\ \text{nm} \times 20\ \text{nm}$) and shows the Si(100) surface following flash annealing at 1200°C . Dimer rows are clearly resolved in this image.

dimer rows forming a 2×1 reconstruction and terrace steps zig-zagging across the surface (the preferred terrace step is oriented along [110]). An STM image of this surface is shown in figure 7(c). This means that the formation of registration marks which requires coating the sample with resist is compatible with the formation of clean Si surfaces.

Our results have shown that it is possible to manipulate C_{60} molecules on Si(111) 7×7 and form nanometre scale clusters. In addition we have shown that C_{60} can act to chemically passivate Si(111) 7×7 . We are optimistic that these results, in conjunction with our alignment techniques, will be developed further to fabricate experimental structures with an active region on an atomic or molecular scale.

Acknowledgments

This work was supported by the UK EPSRC. We are grateful to Dr S Thoms (University of Glasgow) for help with electron beam lithography.

References

- [1] Becker R S, Golovchenko J A and Schwartentruber B A 1987 *Nature* **325** 419
- [2] Eigler D M and Schweizer L K 1990 *Nature* **344** 524
- [3] Crommie M F, Lutz C P and Eigler D 1993 *Science* **262** 218
- [4] Becker R S, Chabal Y J, Higashi G S and Becker A J 1990 *Phys. Rev. Lett.* **65** 1917
- [5] Lyding J W, Shen T-C, Hubacek J S, Tucker J R and Abeln G C 1994 *Appl. Phys. Lett.* **64** 2010
- [6] Dagata J A, Schneir J, Harary H H, Evans C J, Postek M T and Benett J 1990 *Appl. Phys. Lett.* **56** 2001
- [7] Snow E S and Campbell P N 1994 *Appl. Phys. Lett.* **64** 1932
- [8] Li Y Z, Chander M, Patrin J C, Weaver J H, Chibante L P F and Smalley R E 1992 *Phys. Rev. B* **45** 13937
- [9] Devel M, Girard C, Joachim C and Martin O J F 1995 *Appl. Surf. Sci.* **87/88** 390
- [10] Viitanen J 1993 *J. Vac. Sci. Technol. B* **11** 115
- [11] Jung T A, Schlitter R R, Gimzewski J K, Tang H and Joachim C 1996 *Science* **271** 181
- [12] STM, electronics and software purchased from W A Technology, UK
- [13] Maruno S, Inanaga K and Isu T 1993 *Appl. Phys. Lett.* **63** 1339
- [14] Beton P H, Dunn A W and Moriarty P 1995 *Appl. Phys. Lett.* **67** 1075
- [15] Dunn A W, Beton P H and Moriarty P 1996 *J. Vac. Sci. Technol. B* **14** 1596
- [16] Xu H, Chen D M and Creager W N 1993 *Phys. Rev. Lett.* **70** 1850
- [17] Chen D, Chen J and Sarid D 1994 *Phys. Rev. B* **50** 10905
- [18] Hong H, McMahon W E, Zschack P, Lin D-S, Aburano R D, Chen H and Chiang T-C 1992 *Appl. Phys. Lett.* **61** 3127
- [19] Hamza A V, Dykes J, Moseley W D, Dinh L and Balooch M 1994 *Surf. Sci.* **318** 368
- [20] Rosolen G C, Hoole A C F, Welland M E and Broers A N 1994 *Appl. Phys. Lett.* **64** 1932

High spatial resolution spectroscopy of single semiconductor nanostructures

T D Harris, D Gershoni, L Pfeiffer, M Nirmal, J K Trautman and J J Macklin

Lucent Technologies, Bell Laboratories, Murray Hill, NJ 07974, USA

Abstract. Low-temperature near-field scanning optical microscopy is used for the first time in spectroscopic studies of single, nanometre dimension, cleaved edge overgrown quantum wires. A direct experimental comparison between a two-dimensional system and a single genuinely one-dimensional quantum wire system, inaccessible to conventional far-field optical spectroscopy, is enabled by the enhanced spatial resolution. We show that the photoluminescence of a single quantum wire is easily distinguished from that of the surrounding quantum well. Emission from localized centres is shown to dominate the photoluminescence from both wires and wells at low temperatures. A factor of three oscillator strength enhancement for these wires compared with the wells is concluded from the photoluminescence excitation data. We also report room-temperature spectroscopy and dynamics of single CdSe nanocrystals. Photochemistry, trap dynamics and spectroscopy are easily determined.

1. Introduction

Modern epitaxial growth techniques allow thickness control of one molecular layer. This dimensional precision has enabled a wide variety of fundamental and applied physics studies. Optoelectronics based on very thin layers of semiconductor heterostructures, such as quantum wells (QWs), in which carriers are confined to two dimensions (2D), are now dominant for many commercial applications. In an attempt to gain further from the reduction of dimensionality, a world-wide research effort to bring 1D quantum structures such as quantum wires (QWRs) and 0D quantum structures such as quantum dots to the same degree of perfection achieved in the 2D quantum systems has been under way during the last decade [1]. The principal barrier to progress is achievement of adequate uniformity in the second- or third-dimensional confinements. While there is no substitute for dimensional uniformity, we show that restriction of the measurements to single quantum wires or single quantum dots permits understanding of the quantum physics not possible for study of ensembles. We discuss in turn low-temperature near-field scanning optical microscopy (NSOM) of single GaAs quantum wires, and room-temperature confocal scanning optical microscopy of chemically synthesized CdSe quantum dots.

2. Single GaAs quantum wires

Among the most promising ways to achieve dimensional uniformity in 1D structures is cleaved edge overgrowth (CEO) [2]. This technique utilizes two orthogonal

directions of epitaxial growth, exploiting the precision of layer thickness control to form uniform intersecting planes of semiconductor. Two different CEO quantum wire systems have been fabricated and studied:

(i) Strained layer QWRs (SQWRs) in which confinement to 1D is produced by one-dimensional pseudomorphic strain [3]. This strain is induced in the (110)-oriented cleaved edge QW by a (100)-oriented strained layer QW.

(ii) T-shaped QWRs (TQWRs) in which quantum confinement to 1D is produced along the intersection line between the planes of a (100)-oriented QW and that of a (110)-oriented cleaved edge overgrown QW [4]. The lateral dimensions of single QWRs produced by this technique are somewhat uncertain, but comparable to the dimensions of the intersecting QW layers from which they are formed.

Since both wire structures are formed by perturbation of quantum wells, these wires are surrounded by the QW from which they are formed. Study of single CEO QWRs requires care to ensure the unambiguous separation of wire and well spectroscopy. Far-field optical spectroscopy, which has been the most common tool for the characterization of 2D structures, suffers a substantial obstacle [5]. The probed volume of a QW is orders of magnitude larger than the probed volume of a QWR. The QWR spectral features are probably obscured by or attributed to QW features. We show here that the enhanced spatial resolution of low-temperature near-field scanning optical microscopy (LT-NSOM) surmounts this obstacle, permitting unambiguous single QWR studies. There is substantial existing

spectroscopic data on multiple and arrays of CEO QWRs [3,4,6–8]. The added clarity of probing a single wire structure is considerable, since heterogeneity and carrier and electromagnetic field interactions between neighbouring wires are eliminated. We report here low-temperature near-field imaging spectroscopy of the first CEO system, SQWRs.

3. Experiment

The low-temperature NSOM microscope used for this study is described in detail elsewhere [9]. This system was used previously for studies of multiple ‘T’ quantum wires. We describe here only modifications made to the system in order to allow simultaneous excitation and collection in the near field, required for the study of single wires [10]. All the data reported here used excitation and collection with the same aluminium-coated, tapered fibre probe at a sample temperature of 4 K. The excitation radiation was launched to the fibre from a computer-controlled Ti:sapphire laser. To minimize both mode and amplitude intensity fluctuations and permit polarization control, an intensity stabilizer (‘noise eater’) and quarter- and half-wave plates were inserted between two separate sections of the Corning 850 nm single-mode fibre used for our NSOM probes. Fusion spliced into the second section of the fibre is a 3dB two-way splitter. It provides a convenient means of simultaneously monitoring the light sent to the probe, and detecting the light collected by the probe. In order to excite and detect through the same tip, efficient discrimination between excitation and emitted light is required. Elastically scattered light is efficiently filtered by a triple spectrometer. The fibre used in this study had no significant fibre fluorescence for the relevant excitation wavelength (700–800 nm). We thus attribute the background light to Raman scattering within the fibre itself. This background light is highly structured and typically one to two orders of magnitude larger than the photoluminescence (PL) signal collected by the tip (diameter > 200 nm) and strongly polarized parallel to the exciting radiation. By careful adjustment of retardation plates in both the excitation and detection channels, extinction ratios of sample luminescence to background of as much as three orders of magnitude could be achieved. This extinction was stable in time and it allows for reliable discrimination against the background. The light emitted around the fibre tip was collected by a reflecting objective contained within the cryostat. The collimated light from this objective was directed to a second monochromator equipped with a CCD camera, allowing simultaneous monitoring of the entire emission spectrum from the sample in both the near and far fields.

The CEO sample used for this study is schematically described in figure 1. In the first growth step, molecular beam epitaxy (MBE) was used to grow five strained $\text{In}_{0.10}\text{Ga}_{0.90}\text{As}$ QWs of 300, 150, 75, 38 and 18 Å respectively, on a (100)-oriented GaAs substrate. The strained QWs, are separated by 1.0 μm thick layers of GaAs and capped by a 2 μm thick GaAs layer. The sample was then thinned, scribed and mounted in a second

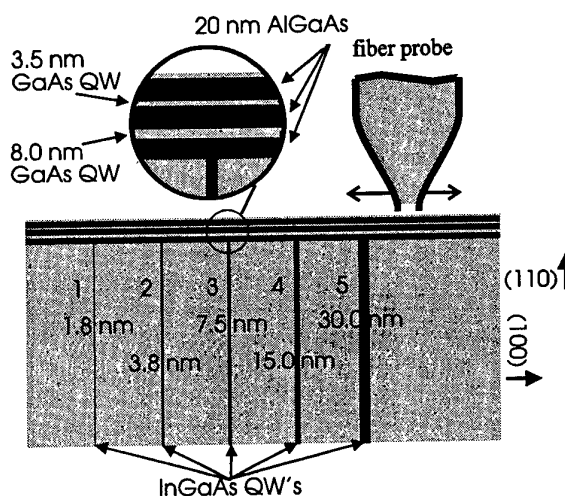


Figure 1. Schematic diagram of the CEO sample structure. The (100) InGaAs QW separation is 1.0 μm , with the fibre tip drawn to this scale to illustrate the expected resolution.

MBE machine. It was cleaved during growth, thus exposing a (110)-facet, onto which the following layers were grown in succession: 200 Å $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, 80 Å GaAs, 200 Å $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, 35 Å GaAs, 200 Å $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ and 50 Å GaAs. On the original substrate (110) face are thus formed 10 single SQWRs, one wire in each of two (110) QWs, for each of the five (100) strained InGaAs QWs. Low-temperature cathodoluminescence (CL) was used previously to show that emission from the AlGaAs/GaAs (110) QWs is red shifted directly above the strained layers. Here, using NSOM spectroscopy, we overcome two major disadvantages of CL: (i) the loss of spatial resolution due to the large diffusion length ($\sim 1 \mu\text{m}$) of the high excess energy cathodo-excited carriers and (ii) the lack of excitation energy tunability.

Two modes of data acquisition were used: PL spectral images are generated by fixing the excitation energy and recording a PL spectrum for 1–5 s at each tip position. A large four-dimensional (x , y , λ , I) data set is thus generated. Alternatively, at a single tip position (image pixel $-x$, y), emission intensity integrated over a selected spectral range is recorded as a function of the excitation energy.

4. Results

In figure 2(a), we show four PL spectra which were selected from the 441 near-field spectra generated in a 21×21 pixel scan of a $2.5 \times 2.5 \mu\text{m}$ square region of the CEO SQWR sample. The spectra are vertically displaced for clarity. In each spectrum a sharp spectral line is observed. These lines arise from carrier recombination within the (100)-oriented strained InGaAs QWs. The spectrum marked ‘1’ results from recombination in the 18 Å strained QW and the lines marked 2, 3 and 4 result from recombination within the 38, 75 and 150 Å strained QWs respectively. The PL line from the 300 Å strained QW is

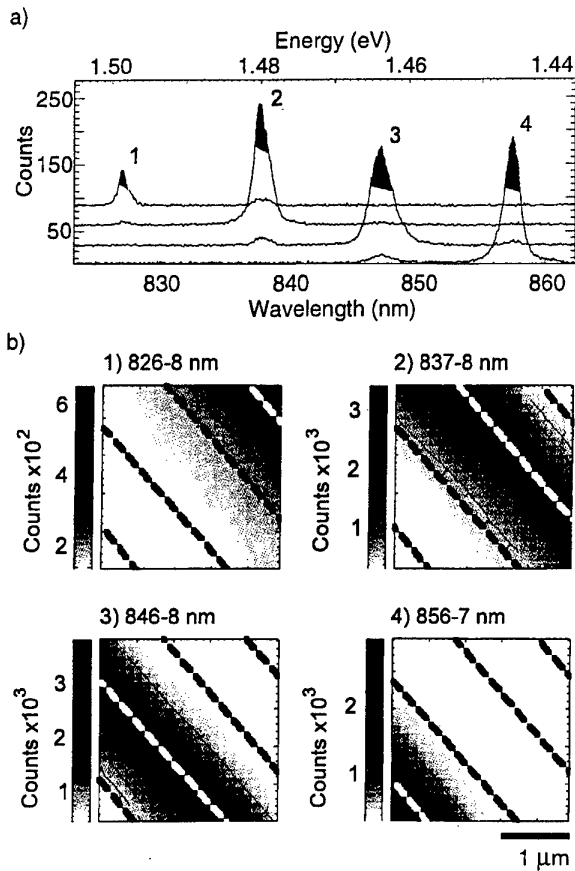


Figure 2. (a) Single-pixel PL spectra from an NSOM image. If the emission intensity is integrated over the shaded wavelength interval the four fixed wavelength emission images shown in (b) result. Images are numbered from the spectra indicating the wavelength interval. Broken lines are drawn to mark the position of the InGaAs strained QWs and transferred to figure 3(b).

not seen in figure 2 since this QW was outside the spatial range of this image.

In figure 2(b) we show the four selective wavelength PL images associated with the spectral lines of figure 2(a). The images are obtained by integrating the PL emission over the wavelength interval marked in gray on figure 2(a). The images are overlaid by a contour plot where each full line represents a 10% change in the emission intensity. The images are numbered in accordance with the spectral domains from figure 2(a). These images allow an accurate determination of the spatial source for each spectral feature. The position of the (100)-oriented strained 18, 38, 75 and 150 Å InGaAs QWs is clearly identified by these images. We have marked these positions by the bold broken lines on the images, for later reference.

In figures 3(a) and 3(b) we show PL spectra and fixed wavelength images for recombination associated with the 80 Å (110)-oriented GaAs/AlGaAs QW for the same scan area as for figure 2. The four broken bold lines of figure 3(b) mark the position of the strained (100)-oriented InGaAs QWs as determined from figure 2(b). The uppermost spectrum in figure 3(a) is dominated by a

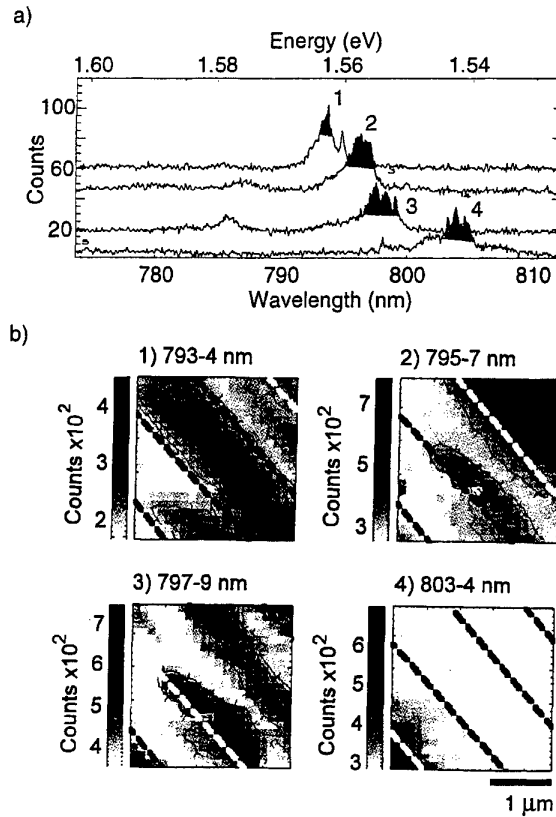


Figure 3. (a) Single-pixel PL spectra from the same area as figure 2(b) for a wavelength near the 80 Å (110) QW. (b) Fixed-wavelength images for the shaded intervals shown above.

spectral line which peaks at 793 nm. This PL emission line is typical of the (110)-oriented CEO 80 Å GaAs/AlGaAs QW as verified by far-field spectroscopy in this work and previous studies [3, 7, 8]. This spectrum is from a pixel midway between the 38 and 75 Å strained QW. The image clearly shows that the spatial origin of this emission strongly anticorrelates with the positions of the strained QWs, peaking between them. The lower three spectra in figure 3(a) originate from pixels above the three strained QWs. The spectral line marked 2 arises above the 18 Å strained QW and the lines marked 3 and 4 originate above the 75 and 150 Å QWs respectively. The fixed wavelength images of these spectral lines strongly correlate with the spatial position of the (100)-oriented strained InGaAs QWs, as can be seen in figure 3(b). There can be no doubt that the strain field of the underlying (100) InGaAs layers perturbs the (110) GaAs QW and shifts its emission to lower energy relative to the emission from the unperturbed QW. This perturbation affects the semiconductor band structure only directly above the strained InGaAs QWs, and thus along a very well defined direction in the (110)-oriented QW plane. This direction defines the SQWR where carriers are confined to narrow stripes within the (110)-oriented GaAs QW plane. The three lower PL spectra in figure 3(a) are thus assigned to carrier recombination within the single SQWRs. The magnitude of these shifts, as large as 20 meV for the 150 Å SQWR from both the 80 and 35 Å (110)-

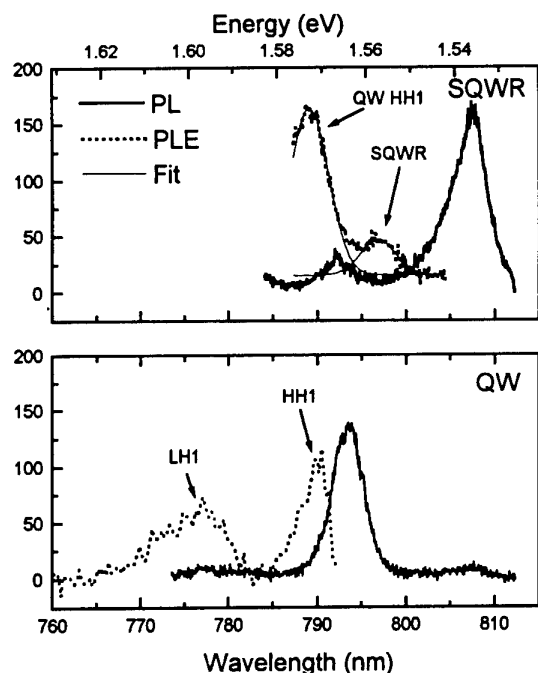


Figure 4. Near-field PL (full curve) and PLE (dotted curve) spectra of the 80 Å (110) QW and the 150 Å × 80 Å SQWR. A two-Gaussians model fit used to determine the relative intensity is shown by the broken line.

oriented QWs (the latter is not discussed in this study), is in agreement with previous measurements using far-field optics and a large array of QWRs [3,7,8]. As expected, the magnitude of the shift scales with the width of the strained layer which determines the dimension of the lateral confinement [11]. Another demonstration of the confinement to 1D can be readily seen in the images of figure 3(b) and the spectra of figure 3(a). We note that both the PL images and the near-field PL spectra (sharp spectral 'spikes') clearly indicate that the emission from both the QWs and QWRs originate from fully (0D) localized centres [12]. It is clearly seen that the near-field PL spectra from the QWRs are broader and have a greater number of sharp spectral lines than that of the QW. This difference obtains from the limitation of carrier diffusion in one of the lateral dimensions.

In figure 4 we show the PL (full curve) and PLE (broken curve) spectra of the 150 Å SQWR and the (110)-oriented 80 Å GaAs QW from a position between the strained (100)-oriented InGaAs QWs. As discussed above, both the excitation and collection are through the NSOM tip. The all near-field PLE spectrum of the (110) CEO GaAs QW is indistinguishable from the far-field PLE spectrum (not shown). We use the similarities to verify that background light originating in the fibre was correctly subtracted. The excitonic transitions associated with the first heavy-hole excitons (HH1) and the first light-hole exciton (LH1) are marked in figure 4. The 5 meV Stokes shift between the HH1 transition measured in PLE and that measured in PL is typical of (110)-oriented QWs [13]. Since the



Figure 5. Photoluminescence emission image of 4 nm CdSe crystallites excited at 532 nm. Resolution is determined by the 1.4 NA oil immersion objective used in this scanning confocal microscopy image. Detection is with a silicon avalanche photodiode.

NSOM tip, 3000 Å diameter, is far larger than the SQWR, carrier diffusion from the well to the wire cannot be avoided. Consequently all the spectral features associated with QW absorption are seen in the PLE spectrum of the SQWR. The spectral feature centred at 799 nm in the PLE spectrum of the SQWR is at a lower energy than the QW band edge and thus can only be assigned to SQWR absorption. With a few reasonable assumptions, the data of figure 4 can be used to experimentally test for the first time the prediction of enhanced oscillator strength for QWR structures [14]. We assume that all photogenerated carriers up to a tip radius diffuse to recombine in the SQWR. This assumption is supported by the spectra and images of figure 3. There is negligible emission from the QW when the tip is positioned directly over a QWR. Since the magnitudes of PL emission from the SQWRs and the QW are comparable, non-radiative recombination can be reasonably ignored. Thus the ratio of QW and QWR oscillator strength can be measured by comparing the peak area of the lowest energy transition of the QW to that of the SQWR as observed in the PLE spectrum of the SQWR, correcting for the geometric area of the two structures. The QW to SQWR area ratio under the tip is approximately 20:1, while the QW-QWR PLE intensity ratio is roughly 6:1 as determined by the two-Gaussians model fit to the data shown in figure 4. A factor of three enhancement in the absorption of the wire with respect to that of the well is thus determined. This demonstration of the increase in the QWR oscillator strength is in agreement with theoretical estimations [14].

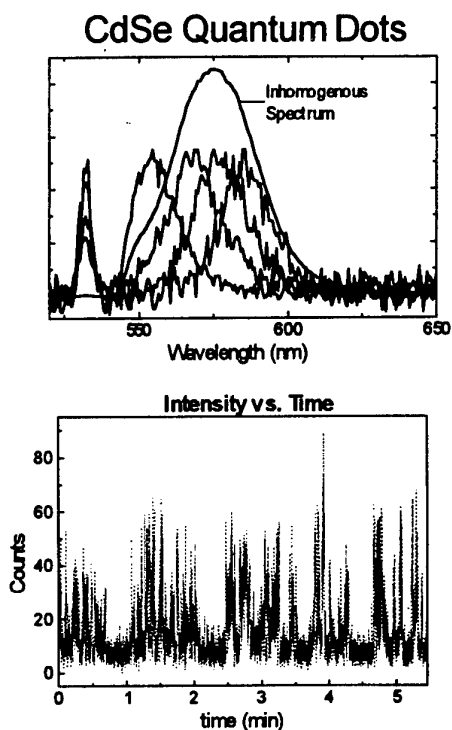


Figure 6. Upper panel: photoluminescence spectra of several particles compared with a spectrum of a large number. Lower panel: sequential time record of photon arrival times for one particle in the image of figure 5. Non-exponential statistics characterize the dynamics at both long and short times.

5. CdSe single quantum dots

Recently, single-molecule characterization has been revolutionized by NSOM and CSOM (confocal scanning optical microscopy) [15–17]. We show here recent room-temperature studies of chemically synthesized CdSe nanocrystals [18]. Heterogeneity has been a particularly difficult barrier to understanding the photophysics of nanocrystals. In figure 5 we show an image of 3.5 nm CdSe nanocrystals dispersed onto a polymer-coated silica cover-slip. Since lateral resolution is $\sim 0.3 \mu\text{m}$, the diffraction limit, coverage is adjusted to give appropriate particle separation. Details of the instrumentation were described previously as were sample preparation procedures [17, 18]. We show here preliminary spectroscopy and dynamics to illustrate the advantage of this experimental strategy. In figure 6 we show representative spectra and photoemission dynamics of one CdSe particle. The PL spectrum line width is typically half the ensemble line width for the best available samples. Even more dramatic is the shift to high energy with time. The most immediate interpretation is that the particle is undergoing photochemistry, probably oxidation, yielding a smaller particle with time. The photochemistry is also reflected in the emission dynamics. The emission time record shows alteration of the trap dynamics. The emission decay time is typically 20 ns. However, there is clear evidence of a long-lived trap state which results in millisecond long 'dark' periods. The nature of the

traps is currently under study. Variation of temperature and excitation density will permit determination of both trap energetics and kinetics of state crossing.

This measurement strategy allows unambiguous determination of the effect of surface passivation chemistry. Differentiation of trapping probability versus trap depth and lifetime is greatly simplified by single-particle studies. The observation of simultaneous spectral shifts, not visible for ensembles, and increased trapping times reveals that chemical evolution is the primary avenue for particle stabilization. Repetition of the measurement in inert atmospheres will immediately reveal the role of internal chemical stability.

6. Conclusion

Using NSOM spectroscopy, we produced photoluminescence and photoluminescence excitation spectra of single quantum wires for the first time. The spatial position of strained, cleaved edge overgrown (110) quantum wires is coincident with the underlying strained (100) quantum wells, and the magnitude of the SQWR energy shift scales with this (100) QW width. From the near-field PLE spectrum we estimate a factor of three oscillator strength enhancement for this semiconductor quantum wire relative to a comparable QW. Using room-temperature scanning confocal microscopy we are able to study the excited state decay dynamics of single nanometre dimension CdSe crystallites. Unique insight is gained for measurements of single entities compared to ensembles.

References

- [1] Arakawa Y and Sakaki H 1982 *Appl. Phys. Lett.* **40** 939
- [2] Arakawa Y, Vahala K and Yariv A 1984 *Appl. Phys. Lett.* **45** 950
- [3] Pfeiffer L N, West K W, Stormer H L, Eisenstein J, Baldwin K W, Gershoni D and Spector J 1990 *Appl. Phys. Lett.* **56** 1697
- [4] Gershoni D, Weiner J S, Chu S N G, Baraff G A, Vandenberg J M, Pfeiffer L N, West K W, Logan R A and Tanbun-Ek T 1990 *Phys. Rev. Lett.* **65** 1631
- [5] Goni A R, Pfeiffer L N, West K W, Pinzuck A, Baranger H U and Stormer H L 1992 *Appl. Phys. Lett.* **61** 1956
- [6] Dingle R, Wiegmann W and Henry C H 1974 *Phys. Rev. Lett.* **33** 827
- [7] Wegscheider W, Pfeiffer L N, Dignam M M, Pinzuck A, West K W, McCall S L and Hull R 1993 *Phys. Rev. Lett.* **71** 4071
- [8] Gershoni D, Katz M, Wegscheider W, Pfeiffer L N, Logan R A and West K W 1994 *Phys. Rev. B* **50** 8930
- [9] Gershoni D, Weiner J S, Fitzgerald E A, Pfeiffer L N and Chand N 1993 *Optical Phenomena in Semiconductor Structures of Reduced Dimensions* ed D J Lockwood and A Pinzuck (Dordrecht: Kluwer) p 337
- [10] Grober R D, Harris T D, Trautman J K and Betzig E 1994 *Rev. Sci. Instrum.* **65** 626
- [11] Grober R D, Harris T D, Harris J K, Betzig E, Wegscheider W, Pfeiffer L and West K 1994 *Appl. Phys. Lett.* **64** 1421
- [12] Baraff G A and Gershoni D 1991 *Phys. Rev. B* **43** 4011
- [13] Hess H F, Betzig E, Harris T D, Pfeiffer L N and West K W 1994 *Science* **264** 1740
- [14] Gershoni D, Brenner I, Baraff G A, Chu S N G, Pfeiffer L N and West K 1991 *Phys. Rev. B* **44** 1931

- [14] Suemune I and Coldren L A 1988 *IEEE J. Quantum Electron.* **24** 1778
- [15] Betzig E and Chichester R 1993 *Science* **262** 1422
- [16] Trautman J K, Macklin J J, Brus L E and Betzig E 1994 *Nature* **369** 40
- [17] Macklin J J, Trautman J K, Harris T D and Brus L E 1996 *Science* **272** 255
- [18] Murray C B, Norris D J and Bawendi M G 1993 *J. Am. Chem. Soc.* **115** 8706

Growth and transport properties of thin Bi films on InP(110)

B G Briner^{†||}, R M Feenstra^{‡||}, T P Chin[§] and J M Woodall[§]

[†] Abt. Oberflächenphysik, Fritz-Haber-Institut, Faradayweg 4–6, 14195 Berlin, Germany

[‡] Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[§] School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA

Abstract. Growth and lateral charge transport properties of thin ($d = 20\text{--}30\text{ \AA}$) Bi films are investigated with scanning tunnelling microscopy. Bismuth is deposited at $T = 140\text{ K}$ onto the cleaved (110) surface of an InP-based heterostructure. Growth at low temperature is kinetically limited and leads to strained, metastable overlayers. After annealing to 300 K the Bi surface consists of atomically flat terraces separated by 12 \AA deep holes. We find that prolonged injection of a high lateral current promotes significant changes in surface morphology which are attributed to a strain relaxation process mediated by electromigration. Scanning tunnelling potentiometry is applied to probe the local response of the semimetal overlayers to the injected lateral current. The observed potential distribution provides evidence for both phonon and defect scattering. At the position of holes and grain boundaries we find typically 2–4 mV high potential steps. It is argued that these steps indeed reflect a localized increase of the film resistance and cannot be attributed to tip-convolution artefacts.

1. Introduction

The scanning tunnelling microscope (STM) has evolved into a powerful and flexible experimental tool which provides insights into a broad variety of atomic-scale processes on metal and semiconductor surfaces [1]. STM-based experiments aimed at investigating charge transport in low-dimensional structures fall into two different categories. Studies on *perpendicular* transport across small barriers, clusters or constrictions use the tunnelling current as the source for the transport process. The excellent spatial confinement of the tunnelling current makes it possible to probe structures down to the atomic scale. Information on the specific interactions which characterize the transport process is reflected in the dependence of the conductivity across the nanostructure on the applied bias voltage. Examples for such experiments are the observation of single-electron tunnelling in double barrier structures [2], and the investigation of quantized conduction across small wires pulled with the STM tip [3]. Thanks to the small size of the investigated structures both studies succeed in finding evidence of quantum transport even at room temperature. In addition, the project on quantum wires nicely demonstrates how the STM can be employed to both fabricate and probe a nanostructure.

The second group of transport-related experiments with an STM concentrates on *lateral* transport in planar structures. Here the source current for the sample under investigation is supplied via external contacts while the STM tip probes the local electrochemical surface potential with almost atomic lateral resolution. This scanning tunnelling potentiometry (STP) technique was introduced by Murali and Pohl to study charging effects in granular Au films [4]. The same authors applied STP for cross-sectional imaging of a GaAs/AlGaAs heterostructure, and illustrated how the shape of the depletion regions depends on the lateral bias voltage [5]. Kirtley *et al* improved the sensitivity of STP to the $10\text{ }\mu\text{V}$ level, and reported the observation of potential steps at grain boundaries of 600 \AA thick granular $\text{Au}_{60}\text{Pd}_{40}$ films [6]. These discontinuities in the potential were ascribed to strong scattering at the grain boundaries, in line with a long-standing theory by Landauer [7] which predicts that localized scatterers give rise to local current and field variations. A subsequent study [8] seriously questioned the interpretation of Kirtley *et al*, and argued that potential steps in STP images can appear as artefacts caused by convolution of the STM tip with the substantial corrugation of the investigated surface. Such tip-convolution effects get larger on a rough sample surface, and they also increase in size for a blunt probe tip. This behaviour is opposite to the intuitive expectation based on topographic imaging where a poor-quality tip produces blurred and featureless results.

^{||} Experiments performed at: IBM, T J Watson Research Center, Yorktown Heights, NY 10598, USA

Several groups have tried to further improve the detection sensitivity of STP [9–11] with the motivation of finding evidence for defect scattering on flat parts of the metal film where tip-convolution artefacts can be excluded. Instead of further optimizing the detection technique we have taken an alternative approach to tunnelling potentiometry which is based on a novel sample geometry and has made it possible to unambiguously observe scattering-induced discontinuities in the surface potential of thin Bi films [12]. Here we describe these experiments in detail with particular emphasis on the growth of the semimetal films and on the observation of potential steps at grain boundaries.

2. Experimental details

For the present experiments our goal has been to optimize the sample geometry in order to overcome the limitations imposed on STP by a rough surface. On the insulating substrates (quartz, mica) which were used in earlier potentiometry studies metal growth usually proceeds via island formation. To obtain continuous films requires the deposition of several hundred angstroms thick layers with obvious consequences for the surface corrugation. In contrast, it has been demonstrated that it is possible to grow continuous and flat semimetal (Bi, Sb) films on the cleaved (110) surface of GaAs and InP [13–16]. To benefit from these favourable growth conditions we have developed a substrate which consists of an n–i–n heterostructure on top of an InP wafer. The heterostructure is used to inject a high lateral current into a narrow segment of an evaporated semimetal film. Figure 1 shows a schematic layout of this heterostructure together with a functional diagram of the STP set-up. The choice of degenerately n-doped $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ for the contact layers is dictated by the need to avoid the formation of Schottky barriers at the interfaces to the evaporated conductor which would inhibit efficient current injection. For a dopant concentration of $n = 10^{19} \text{ cm}^{-3}$ the Fermi level of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ is significantly shifted into the conduction band, and it can be expected to form ohmic contacts to the metal overlayer. The 5000 Å thick spacer layer consists of either semi-insulating, Fe-doped InP or low-temperature MBE-grown, slightly p-doped $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$. This spacer is separated from the p-contact layers by two 1000 Å thick, moderately n-doped $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ buffers. We have found that, in particular for the substrates with $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ spacer layers, these buffers are necessary to avoid interband tunnelling which would result in a low shunt resistance. To provide electrical contacts to the substrate holder, gold films are evaporated onto the topmost layer of the heterostructure and onto the back side of the InP-wafer. The top contact is left unannealed, because diffusion of Au into the heterostructure would create shunts across the insulating spacer.

To prepare a flat semimetal overlayer we start by cleaving the samples in the same ultra-high vacuum (UHV) system which is used for the STP experiments. Immediately after cleavage, the substrates are cooled down to $T = 140 \text{ K}$ and bismuth is evaporated onto the cleaved edge. Before resorting to this low-temperature deposition technique we

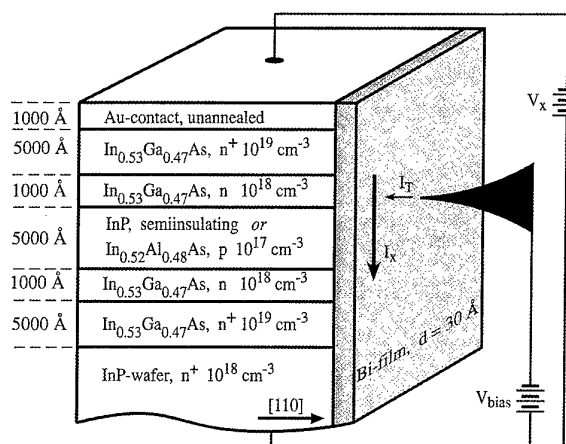


Figure 1. The sample geometry. Bi is deposited onto the cleaved edge of an n–i–n heterostructure. The potential drop on the portion of the Bi film bridging the i layer is probed with the STM.

have investigated the room temperature growth of Bi and Sb. Unfortunately, for both materials it was impossible to obtain flat films at $T = 300 \text{ K}$. During Bi evaporation the sample resistance is continuously monitored to check for the formation of a conducting bridge across the i layer. For deposition at 140 K the resistance starts to drop at a threshold of only 6 Å. This is a clear indication for the flat and continuous growth of the Bi films. Analogous experiments were performed at different temperatures and with Sb and Be as alternative evaporation materials. We found that for these materials a much thicker layer ($>50 \text{ Å}$) had to be deposited before the sample resistance started to decrease. It should be emphasized that due to imperfections in the i layer the percolation limit cannot be properly detected, since at this point the overlayer still has a much poorer conductance than the substrate. Nevertheless, STM topographs indicate a clear correspondence between a rough surface morphology and a large percolation threshold during growth. Because all STP experiments are performed at room temperature we have to anneal the Bi films to 300 K after low-temperature deposition. We find that the thinnest Bi films become discontinuous upon heating to 300 K. Therefore we are presently restricted to the investigation of films with a thickness of at least 15–20 Å. All layer thicknesses quoted in this study are derived from the readings of a crystal deposition monitor i.e. they should be understood as average values.

Topographic and potentiometric data on these Bi films are obtained with a UHV STM equipped with facilities for *in situ* tip and sample change [17]. Single-crystal W tips are cleaned by sputtering and heating until they show a sharp and stable field-emission pattern. STM topographs are recorded with a constant tunnelling current of 100 pA and a sample bias of the order of 100 mV. For the potentiometry experiments we disable the STM feedback loop at each image pixel, and record a current–voltage (I – V) curve which typically extends over an interval of $\pm 50 \text{ mV}$ around the local electrochemical potential V_0 on the surface. The tip–surface distance is intentionally

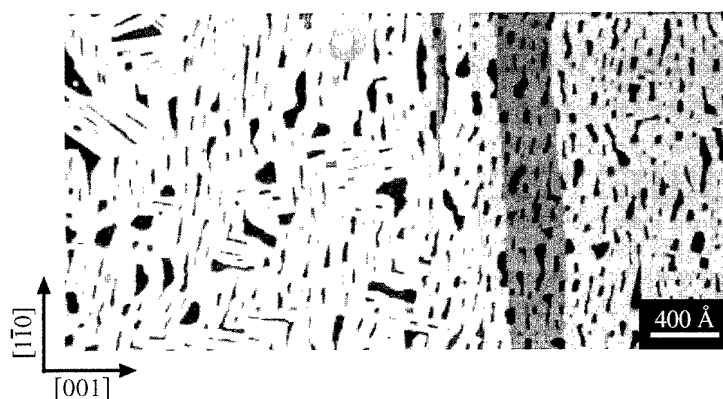


Figure 2. Topography image of a 20 Å thick Bi film grown at $T = 140$ K on InP(110). $V_{bias} = 250$ mV, $I_T = 100$ pA, field of view: $3900 \text{ Å} \times 1950 \text{ Å}$.

reduced by 1 Å after disabling the feedback loop in order to increase the current-detection sensitivity close to the zero-crossing point. After the experiment, the measured $I-V$ data are interpolated with third-order polynomials, and an STP image is generated from the zeros of these fit curves which represent the experimental approximation for V_0 . We find that in our set-up the detection sensitivity is currently limited to 0.5 mV by external noise sources. The spectroscopic data confirm that all investigated Bi films are good conductors because their $I-V$ curves are almost linear close to V_0 .

3. Results and discussion

3.1. Growth of Bi on InP(110)

The investigation of 20–30 Å thick Bi films using an STM shows that under the described growth conditions bismuth forms continuous overlayers which are characterized by atomically flat planes interspersed with irregularly shaped holes. The typical surface morphology found in the present study is illustrated in figure 2, a large-scale STM topograph recorded on a nominally 20 Å thick Bi layer. Many small, always 12 Å deep, holes are discernible. This large step height corresponds to the size of the pseudocubic unit cell of Bi in the (111) direction. The formation of such large steps is somewhat unexpected. Structural studies of thicker Bi layers deposited at room temperature on GaAs [15] and InP [14] found that the surface usually consists of terraces separated by 4 Å high steps. We attribute the morphology of our films to the growth at low temperature which is kinetically limited and consequently leads to strained and metastable overlayers. Patrín *et al* [16] have reported a similar influence of the deposition temperature on the surface structure of Bi films grown on GaAs and InP, but they find smaller (7 Å) steps, probably due to different film thicknesses and growth conditions. The elongated vertical steps visible in figure 2 are cleavage defects on the heterostructure which are replicated in the semimetal overlayer. As will be shown below, this particular growth mode of Bi is a favourable prerequisite for the potentiometry experiments because apart from the holes the film surface is flat on a scale of several 1000 Å.

The hypothesis that after growth the films are in a metastable state is confirmed by the observation that injecting a high lateral current for a prolonged time induces significant changes in surface morphology. Figure 3 shows a $7900 \text{ Å} \times 4800 \text{ Å}$ topography image ($d_{Bi} = 20 \text{ Å}$) recorded on the part of the Bi film over the heterostructure which is exposed to the lateral current. Before taking this image, an average current density of $3.5 \times 10^6 \text{ A cm}^{-2}$ has been run through the film for 8 h. Near the right border of figure 3 the ‘as-grown’ structure is still visible. However, most of the film surface has drastically changed. The 12 Å deep holes have been replaced by large terraces separated by 4 and 8 Å high steps. In the strained phase, the surface defects are preferentially aligned along the (110) direction of the substrate, a fact which underlines that this growth mode is strongly influenced by the misfit between substrate and overlayer lattices. In contrast, the islands in the relaxed phase are more isotropic in shape, and the step orientations reflect the approximately hexagonal symmetry of the Bi surface. Occasionally, deep holes appear also in this relaxed phase. Such holes present both an opportunity and a technical problem for the potentiometry experiments. On one hand, the deep holes constitute dominant scattering centres for the lateral current which allow for a detailed study of the transport-induced potential distribution. On the other hand, the steep steps and the reduced conductivity inside the holes make scanning across such defects difficult and prone to tip-crashes.

It may be argued that the observed changes in surface morphology are caused by locally heating the Bi film with the applied current. To test whether the strain relaxation process is thermally stimulated we have annealed the sample on which figure 3 was recorded to $T = 420$ K for 30 min. By imaging a part of the semimetal film which has *not* been subjected to a lateral current we can make sure that here any observed deviations from the ‘as-grown’ phase must be thermally induced. Figure 4 shows that the surface morphology has also changed in response to thermal annealing, but the resulting structure is significantly different compared to the current-induced effects. The slightly elongated, 12 Å deep holes are still present and the average lateral size of these defects has significantly increased. Periodic fringes (corrugation

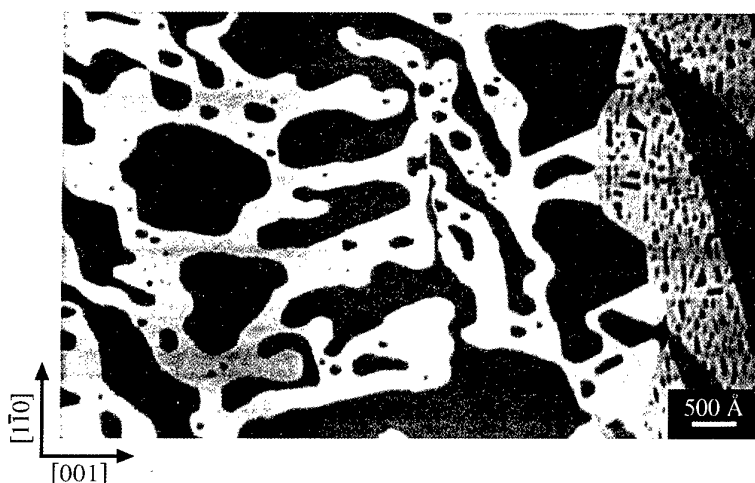


Figure 3. STM topograph ($7900 \text{ \AA} \times 4800 \text{ \AA}$) showing the change of surface morphology in response to lateral current injection. $d_{\text{Bi}} = 20 \text{ \AA}$. The strained phase (still visible near the right border of the image) transforms into large terraces with step heights of 4 and 8 Å.

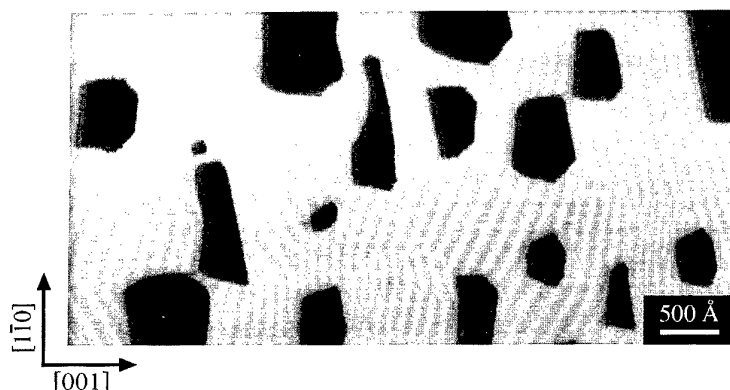


Figure 4. Same film as in figure 3 after annealing to $T = 420 \text{ K}$. Topography image of a part of the Bi film on the InP wafer which has not been subjected to a lateral current. Field of view: $5600 \text{ \AA} \times 2800 \text{ \AA}$.

amplitude $\sim 1 \text{ \AA}$) now appear on the flat parts of the surface. These fringes are similar to the moiré pattern which has been reported by Patrin *et al* [16], and which was explained as a result of the misfit between substrate and overlayer lattice constants. The varying angle of the fringes with respect to the substrate nicely illustrates the existence of differently oriented domains in these polycrystalline films. The observation that domain boundaries often originate from the surface defects lets us suggest that the holes act as pinning centres for the metastable surface. We observe similar fringes with a much reduced corrugation amplitude also on thin ($d = 20 \text{ \AA}$) Bi films in the 'as-grown' phase. Based on the significant difference between figures 3 and 4 we propose that strain relaxation on the part of the film over the heterostructure is promoted by electromigration in response to the high injected lateral current. This interpretation is confirmed by a numerical estimate of Ohmic heating. Assuming intimate contact between film and substrate, we find that a lateral current density of $5 \times 10^6 \text{ A cm}^{-2}$ leads to a temperature rise of less than 1 K. Electromigration alone cannot be made responsible for the current-induced structural changes. It leads to an oriented

motion of atoms along the direction of the current [18], but the surface structure shown in figure 3 does not provide evidence for such an oriented mass flow. Therefore, we tentatively suggest that electromigration sets in at the 12 \AA holes where the local electric field is enhanced (see below). By selectively reshaping and removing the pinning centres, electromigration lowers the potential barrier for thermally assisted strain relaxation which finally transforms the surface into the observed isotropic structure.

3.2. Potentiometry

Potentiometry experiments are carried out on 20–30 Å thick Bi films subjected to current densities of the order of 10^6 A cm^{-2} . They always reveal the coexistence of both phonon and defect scattering in the investigated films. Figure 5 presents simultaneously recorded topography (a) and potential images (b) on a 30 Å thick Bi film. The potential image covers a greyscale range of 34 mV. A horizontal potential ramp is clearly visible. It corresponds to the resistance caused by electron-phonon scattering in a film with constant thickness, and obviously represents

a significant portion of the total film resistance at room temperature. However, in addition to the ramp we find small steps in the surface potential coinciding with the positions of the defects on the topography image. To better illustrate the formation of these typically 2–4 mV high steps, figure 5(c) presents line-cuts of the potential distribution along the direction of the externally applied field. The two holes which cause the depicted potential steps have a horizontal extension a of 30 and 50 Å. The topographic information on the defect size and the measured potential ramp which, for this image, corresponds to a field of $E_x = 31 \mu\text{V Å}^{-1}$ allow us to determine an upper bound for the possible influence of tip-convolution artefacts on the shape of the line-plots shown in figure 5(c). Assuming that no voltage drop occurs at the holes, the distortion of the uniform ramp by tip convolution could at most lead to voltage steps of $\Delta V = E_x a$. As indicated in figure 5(c) these steps are significantly smaller than the observed discontinuities. In contrast to tip convolution which only locally distorts the potential ramp, the steps in the experimental data lead to an offset between the flat parts of the potential curve i.e. they represent a real resistance increase. It must be emphasized that an almost perfect film surface is an important prerequisite for interpreting the potentiometry data. The simple linear background ramp which makes it possible to easily separate phonon and defect scattering is a direct consequence of the fact that apart from the holes the films are atomically flat.

A network of faint lines connecting the surface defects is visible in figure 5(a). These lines are grain boundaries which separate differently oriented crystallites of the Bi film. The potential distribution near a vertical grain boundary has been highlighted in figure 5(b) by locally changing the greyscale repartition. A small step corresponding to $\Delta V_0 < 1 \text{ mV}$ is visible. It must be pointed out that such small potential features are close to the detection limit of our STP set-up which is currently restricted to $\sim 0.5 \text{ mV}$ by external noise sources. We find that for the thinnest studied Bi films the grain boundaries give rise to more pronounced potential discontinuities. A larger misfit-induced strain in these films probably leads to an increased scattering cross-section at the grain boundaries. Figure 6 presents topography and STP images which have been recorded on a 20 Å thick Bi film subjected to a lateral current density of $1.5 \times 10^6 \text{ A cm}^{-2}$. Grain boundaries appear as bright lines in the STM-topograph (a). Due to a somewhat blunt probe-tip the borders of the 12 Å deep holes appear blurred. The potential image (b) displays similar features to those visible in figure 5. However, here the highlighted portion of the image clearly shows the formation of a potential step at the grain boundary. For this film not only the domain boundaries but also the fringe pattern, which has been discussed in the preceding section, can be discerned in both the topography and potential images. This finding illustrates the existence of localized scattering at defects down to the atomic scale.

A rough sample surface does not only increase the chance for tip-convolution, it also makes it difficult to distinguish between artefacts and real potential discontinuities. The background contribution to the

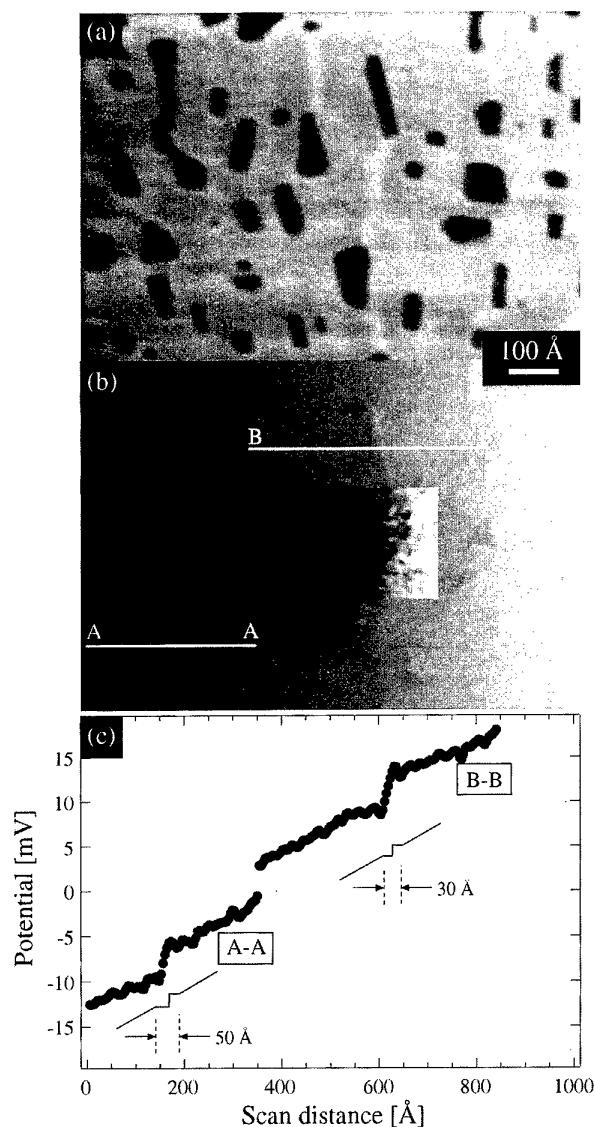


Figure 5. (a) An STM topograph. (b) A simultaneously recorded STP image on a 30 Å thick Bi film. Potential discontinuity at grain boundary is highlighted. (c) This graph presents line-cuts across the potential image as marked in (b) and shows the expected potential slopes if the steps were entirely caused by tip convolution.

resistance of a rough film changes on a small scale, and consequently no smooth potential ramp can be observed by STP. As explained in section 2, the surface quality of the Bi films strongly depends on the deposition temperature. Figure 7(a) shows the surface of a nominally 25 Å thick Bi film which has been prepared by evaporation at a somewhat higher temperature ($T \approx 200 \text{ K}$). The enhanced atomic mobility during deposition obviously has led to the formation of a granular surface structure. On the potential image in figure 7(b) a ramp is still discernible, but the potential distribution is dominated by strong discontinuities which are suggestive of transport across an almost disconnected layer. It is possible that some of these discontinuities appear artificially enhanced. However, tip-convolution cannot be made responsible for the inverted potential steps marked by arrows in figure 7(b). Such a

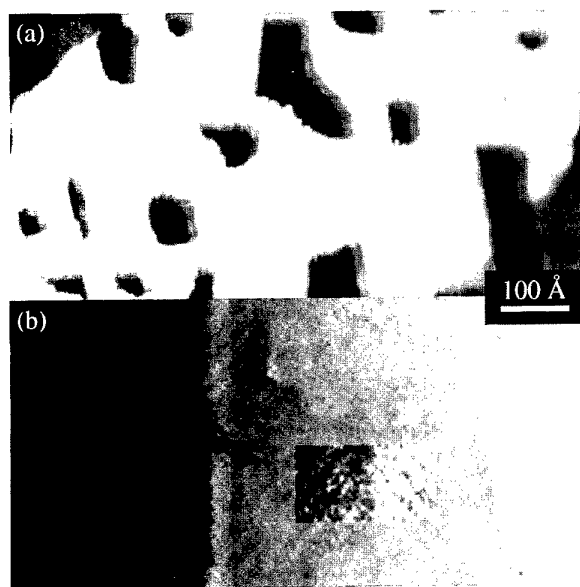


Figure 6. Topography (a) and potential images (b) on a 20 Å thick Bi film. Lateral current density = 1.5×10^6 A cm⁻². A potential step at a grain boundary has been highlighted in (b) by locally changing the greyscale repartition.

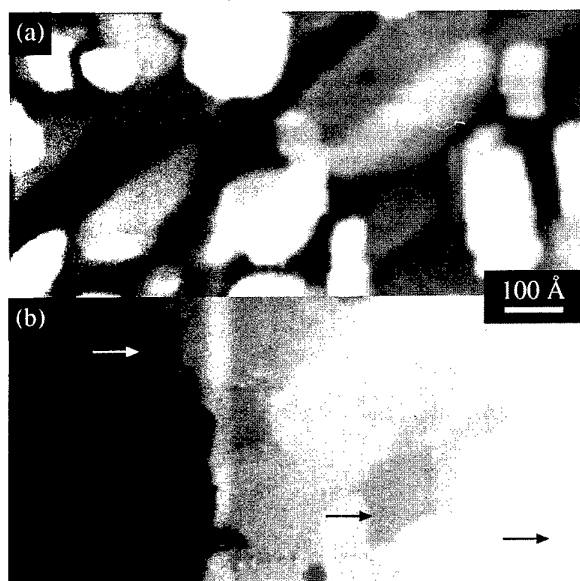


Figure 7. (a) A nominally 25 Å thick Bi film. Growth at $T = 200$ K results in a rough surface. (b) The corresponding potential image is characteristic for an almost disconnected film. Inverted potential steps are marked by arrows.

field inversion indicates either a local current flow opposed to the direction of the applied field or the prevalence of charging effects.

The deep holes which appear on the semimetal films as a result of the described strain relaxation process act as strong scatterers. We have found [12] that they give rise to very pronounced potential discontinuities (step size >20 mV). In the vicinity of a hole, the 'net' potential

after subtraction of a linear background takes on the shape of a two-dimensional dipole. The lobes of the dipole extend over several hundred angstroms beyond the geometric size of the defect. The observation of deviations from the background ramp on *flat* parts of the surface makes it possible to unambiguously assign the dipole to defect scattering even though tip-convolution artefacts *within* the deep holes cannot be excluded. The fact that the observed potential resembles the theoretically predicted resistivity dipole [7, 19] motivates the question about the relative importance of ballistic (electron–electron) and diffusive (electron–phonon) scattering which is discussed in detail in [12]. We find that in the present room-temperature experiments diffusive scattering makes the main contribution to the observed potential discontinuities, but ballistic effects are also identified. From the low carrier density in bulk Bi one would expect that ballistic transport dominates even at room temperature. However, this apparent contradiction is resolved by noting that surface states can significantly increase the effective carrier concentration in thin Bi films [20].

In summary, the presented results provide clear evidence for defect scattering at room temperature i.e. the potential discontinuities represent a direct fingerprint of residual resistivity. The novel sample geometry could also make it possible to observe some of the finer details in the scattering-induced potential distribution, such as the predicted Friedel oscillations [19, 21], if a similar experiment is performed at low temperature.

Acknowledgments

We would like to thank R Landauer and J Kirtley for helpful discussions. One of us (BB) gratefully acknowledges financial support by the Schweizerischer Nationalfonds.

References

- [1] Güntherodt H-J and Wiesendanger R (eds) 1994 *Scanning tunnelling Microscopy I* (Springer Series Surface Science 20) 2nd edn (Berlin: Springer)
- [2] Schönenberger C, van Houten H and Donkersloot H C 1992 *Europhys. Lett.* **20** 249
- [3] Olesen L, Lægsgaard E, Stensgaard I, Besenbacher F, Schiøtz J, Stoltze P, Jacobsen K W and Nørskov J K 1994 *Phys. Rev. Lett.* **72** 2251
Brandbyge M, Schiøtz J, Sørensen M, Stoltze P, Jacobsen K W, Nørskov J K, Olesen L, Lægsgaard E, Stensgaard I and Besenbacher F 1995 *Phys. Rev. B* **52** 8499
- [4] Murali P and Pohl D 1986 *Appl. Phys. Lett.* **48** 514
- [5] Murali P, Meier H, Pohl D W and Salemink H W M 1987 *Appl. Phys. Lett.* **50** 1352
- [6] Kirtley J R, Washburn S and Brady M J 1988 *Phys. Rev. Lett.* **60** 1546
- [7] Landauer R 1957 *IBM J. Res. Dev.* **1** 223
Landauer R 1975 *Z. Phys.* **B 21** 247
- [8] Pelz J P and Koch R H 1990 *Phys. Rev. B* **41** 1212
- [9] Besold J, Reiss G and Hoffmann H 1993 *Appl. Surf. Sci.* **65/66** 23
- [10] Möller R, Baur C, Esslinger A and Kürz P 1991 *J. Vac. Sci. Technol. B* **9** 609
- [11] Koslowski B and Baur C 1995 *J. Appl. Phys.* **77** 28

- [12] Briner B G, Feenstra R M, Chin T P and Woodall J M
1996 *Phys. Rev. B* **54** R5283
- [13] Annovi G, Betti M G, del Pennino U and Mariani C 1990
Phys. Rev. B **41** 11 978
- [14] Resch U, Esser N and Richter W 1991 *Surf. Sci.* **251/252**
621
- [15] Patrin J C, Li Y Z, Chander M and Weaver J H 1992 *Phys.*
Rev. B **46** 10 221
- [16] Patrin J C, Li Y Z, Chander M and Weaver J H 1993
J. Vac. Sci. Technol. A **11** 2073
- [17] Feenstra R M 1994 *Phys. Rev. B* **50** 4561
- [18] See e.g. Landauer R and Woo J W F 1974 *Phys. Rev. B* **10**
1266
Sham L J 1975 *Phys. Rev. B* **12** 3142
- Sorbello R S and Dasgupta B 1977 *Phys. Rev. B* **16** 5193
The relative influence of local electric field (direct force)
and momentum exchange with the moving charges
(electron wind force) on the mass transport is still
debated.
- [19] Zwerger W, Bönig L and Schönhammer K 1991 *Phys. Rev.*
B **41** 6434
- [20] Hoffman C A, Meyer J R, Bartoli F J, Di Venere A, Yi X
J, Hou C L, Wang H C, Ketterson J B and Wong G K
1993 *Phys. Rev. B* **48** 11 431
- Komnik Yu F, Bukhstab E I, Nikitin Yu V and Andrievskii
V V 1971 *Sov. Phys.-JETP* **33** 364
- [21] Sorbello R S 1981 *Phys. Rev. B* **23** 5119

Landau subbands generated by a lateral electrostatic superlattice—chasing the Hofstadter butterfly

Till Schlösser[†], Klaus Ensslin^{†‡}, Jörg P Kotthaus^{†§} and Martin Holland^{||}

[†] Sektion Physik, Ludwig-Maximilians-Universität, D-80539 München, Germany
^{||} Department of Electronics, University of Glasgow, Glasgow G12 8QQ, UK

Abstract. Magnetoresistance studies on two-dimensional electron systems in GaAs–AlGaAs heterojunctions exposed to voltage-tunable square lateral superlattices with periods in the 100 nm range reveal at low temperatures a characteristic splitting of individual Landau bands. Around magnetic field values at which p/q flux quanta penetrate the unit cell of the superlattice, with p and q being small integers, p subbands can be observed. A discussion of which conditions have to be met to reveal such Hofstadter-type splitting of Landau bands illuminates the experimental difficulties that have to be overcome in chasing the Hofstadter butterfly.

1. Introduction

In recent years nanofabrication technologies have enabled us to study ballistic and phase coherent motion of conduction electrons in semiconductors under the action of artificial potential landscapes tailored by the experimentalist. In particular it has become possible to study electronic systems in which geometrically defined length scales well above the crystalline lattice constant but much smaller than typical scattering lengths become comparable to magnetically defined lengths. A widely studied system has been electrons confined to a two-dimensional plane with a magnetic field applied perpendicular to this plane and a one- or two-dimensional periodic electrostatic potential acting on the electrons moving ballistically through this superlattice potential. One motivation for such studies has been the long-predicted quantum commensurability phenomena expected to occur when the unit cell of a square lattice becomes penetrated by a magnetic flux Φ comparable to a flux quantum Φ_0 . For a perturbing magnetic field, Hofstadter has predicted that commensurability between the lattice unit cell and the flux lattice causes a single one-electron band to split up into a fascinating self-similar spectrum of subbands [1]. The opposite case of a strong magnetic field and a weak periodic electrostatic potential yields formally the same spectrum and causes a single Landau band to split up into p subbands whenever the flux Φ per unit cell equals p/q flux quanta with p and q both being integers.

[‡] Present address: Laboratorium für Festkörperphysik, ETH Zürich, CH 8093 Zürich, Switzerland.

[§] E-mail: jorg.kotthaus@physik.uni-muenchen.de.

Whereas the intricate commensurability effects reflected in this so-called Hofstadter butterfly have been the central point of many theoretical studies [2] they have long been experimentally beyond reach, since the magnetic fields which are required to fill a crystalline unit cell with flux comparable to a flux quantum are well above magnetic fields generated in a laboratory. However, with artificial superlattices one is now able to create periodic potentials with periods in the range down to 100 nm in the plane of a high-mobility two-dimensional electron gas (2DEG), generated, for example at a GaAs–GaAlAs heterojunction interface. Hence it is not surprising that a few years ago several groups set out to chase the Hofstadter butterfly. To everybody's surprise, however, the first commensurability phenomenon, found in the magnetoresistance of lateral superlattices by Weiss *et al* [3], in a one-dimensional superlattice occurs whenever the superlattice period becomes commensurate with the classical cyclotron diameter $2R_c$. This effect was correspondingly explained to be of predominantly classical origin [4]. Subsequently also observed in square superlattices [5], we now understand these magnetoresistance oscillations to result from classical ballistic transport across several superlattice periods. Recently we have been able to satisfy the more stringent requirements on the homogeneity and potential amplitude of a small-period square superlattice that are necessary to be able to observe the quantum commensurability exemplified in the Hofstadter butterfly [6]. Here we summarize these studies and demonstrate that we have achieved a first success in chasing the Hofstadter butterfly by observing the occurrence of Landau subbands in accordance with the flux

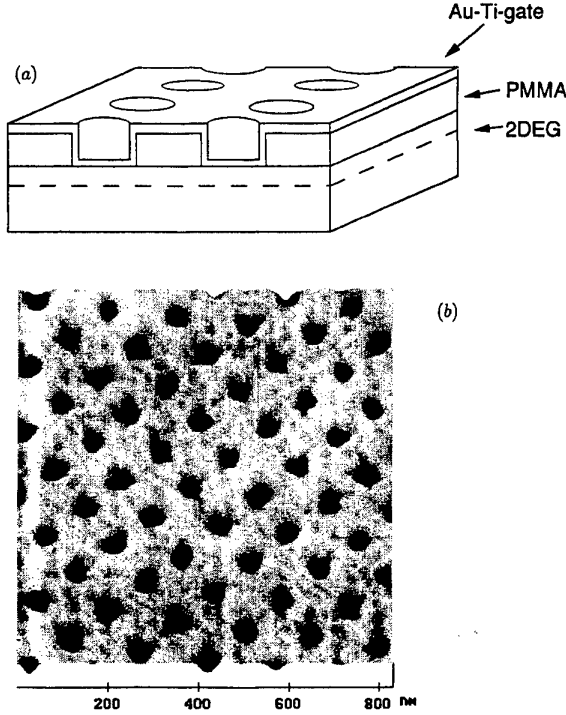


Figure 1. (a) Schematic view of a superlattice sample. (b) AFM image of the patterned electron-sensitive resist layer. The lattice period is 105 nm.

commensurability condition. Our experiments also demonstrate that in real lateral superlattices many-body phenomena, in particular magnetic-field-dependent screening, make the Hofstadter butterfly much more fragile than originally anticipated, thus opening only a relatively small experimental window to the world of quantum commensurability in magnetic fields.

2. Experimental notes

One essential prerequisite for our studies is the preparation of high-mobility 2DEGs exposed to square electrostatic superlattice potentials having periods in the range of 100 nm with tunable potential amplitude. We employ modulation-doped GaAs–AlGaAs heterojunctions grown by molecular beam epitaxy in which the 2DEG is located only 36 nm below the crystal surface [7]. Typically the 2DEGs have carrier densities of $N_s = 5 \times 10^{11} \text{ cm}^{-2}$ and mobilities of $40 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ at $T = 4.2 \text{ K}$ corresponding to a mean free path $l = 4.7 \text{ } \mu\text{m}$ and a Fermi wavelength $\lambda_F = 35 \text{ nm}$. A Hall bar geometry is defined by wet chemical etching with an area of $20 \text{ } \mu\text{m} \times 8 \text{ } \mu\text{m}$ between the voltage probes, i.e. dimensions that are larger than both the elastic and inelastic mean free path. Using electron beam lithography an electron-sensitive resist layer on the sample surface is suitably exposed and developed to create a lattice of voids and then covered by an evaporated metal gate as sketched in figure 1. Due to the proximity of the 2DEG to the sample surface we have the unique possibility of inducing lateral superlattices with periods as small as 80 nm and relatively large and widely tunable potential modulation. We have

Landau subbands generated by an electrostatic superlattice

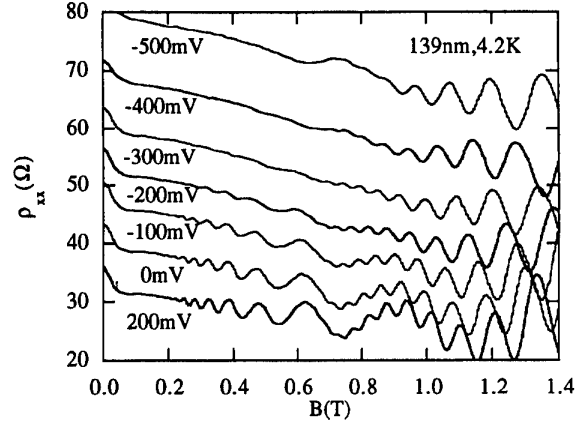


Figure 2. Magnetoresistance traces of a 139 nm period superlattice heterostructure at a temperature of 4.2 K; with potential amplitude increasing from the lowest to the upper curves the amplitude of the classical commensurability oscillations is found to decrease.

studied square superlattice devices with periods of 215 nm, 139 nm, 105 nm and 80 nm. The magnetoresistance measurements are carried out in a dilution refrigerator at a bath temperature of 30 mK and the current sent through the samples is below 10 nA.

3. Magnetoresistance oscillations and commensurability phenomena

A typical magnetoresistance trace at $T = 4.2 \text{ K}$ in a superlattice with period $a = 139 \text{ nm}$ and with varying potential modulation is displayed in figure 2. At gate voltages around 0 V we observe two sets of magnetoresistance oscillations both being periodic in $1/B$. At high magnetic fields above about 0.7 T we observe the well known Shubnikov–de Haas (SdH) oscillations of a 2DEG which reflect the intersection of Landau levels with the Fermi level. At lower magnetic field we observe commensurability oscillations with minima at

$$2R_c = (n - \frac{1}{4})a. \quad (1)$$

In lowest-order perturbation theory the electrostatic potential modulation causes the Landau levels of the 2DEG to transform into Landau bands with periodically oscillating band width whenever the classical cyclotron diameter equals a multiple of the superlattice period. The Landau level bandwidth at the Fermi energy vanishes at the so-called flat-band condition, equation (1). Classically, for example with a 1D potential modulation along the x direction, the electric field E corresponding to the potential modulation causes an $E \times B$ drift along y that reflects the dispersion dE/dk_x of the Landau band at the Fermi energy [4, 8]. This yields an oscillating band contribution to the conductivity σ_{yy} that in lowest approximation becomes superimposed onto the scattering contribution to the conductivity that reflects the density of states at the Fermi energy and, in a sufficiently strong magnetic field where $\mu B > 1$ and at sufficiently low

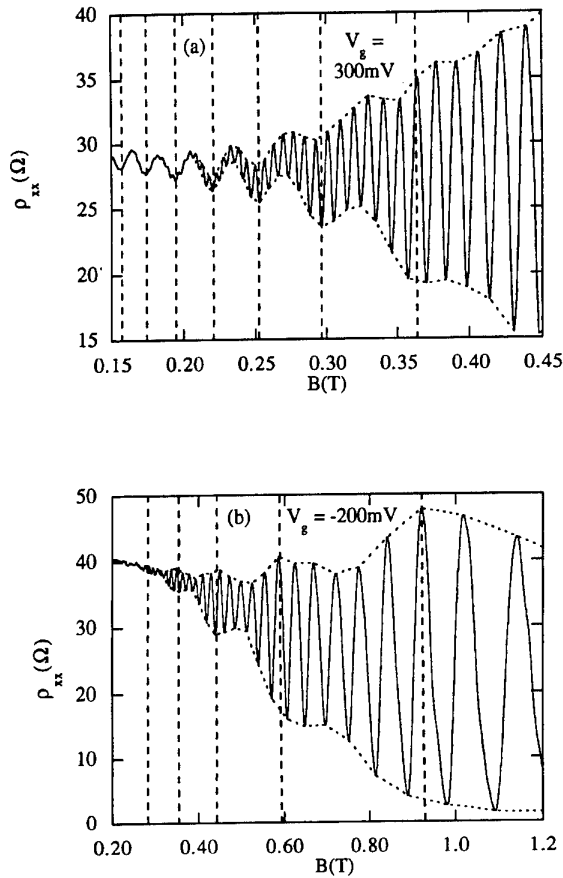


Figure 3. Magnetoresistance traces at a bath temperature of 30 mK. The vertical broken lines indicate magnetic field positions where the Landau bands are essentially flat. (a) At those values and weak potential modulation the experiment reveals minima in the resistance modulation and especially pronounced SdH oscillations as visualized by the envelope of the minima and maxima of the SdH oscillations. In this situation band conductivity dominates the upper SdH envelope. (b) In the regime of larger potential modulation where the Landau bands become relatively wide the upper envelope of the SdH oscillations clearly shows maxima at flat band indicating the breakdown of band conductivity. (From [6].)

temperatures, leads to the SdH oscillations. It should be noted that in this high-field limit the conductivity σ_{yy} becomes roughly proportional to the magnetoresistance ρ_{xx} . According to the picture of classical commensurability one expects the band contribution to the magnetoresistance to increase quadratically with the superlattice potential [4, 8]. Instead we observe in figure 2 that increasing potential modulation decreases the amplitude of the commensurability oscillations. This is a first indication that the classical picture is no longer appropriate.

This becomes even more obvious when we decrease temperature and correspondingly increase thermal resolution and phase coherence length. Then we can observe SdH oscillations at magnetic fields where classical commensurability effects are expected. At first glance figure 3 shows us that classical commensurability oscillations modulate the amplitude of the SdH oscillations, as apparent for the weak

modulation case displayed in figure 3(a). There the band conductivity contribution dominates the upper SdH envelope and produces minima at the flat band condition, equation (1), as indicated by the vertical broken lines. However, at stronger modulations, as in figure 3(b), we observe a distinctly different behaviour. Whereas here the envelope of the SdH minima still exhibits minima at flatband the envelope of the SdH maxima shows maxima at flatband resulting from an increased density of states at the Fermi energy and a correspondingly stronger influence of the scattering-induced contribution to the conductivity. The suppression of the band conductivity contribution in this case indirectly reflects the break-up of the Landau bands into subbands [8].

4. Manifestation of Landau subbands

To be able to observe Landau subbands directly quite a few conditions have to be met. Apart from low temperatures to guarantee a sharp Fermi distribution function, the scattering-induced broadening \hbar/τ as well as broadening caused by spatial inhomogeneities need to be smaller than the expected splitting into Landau subbands. A natural experimental measure of all relevant broadening mechanisms is given by the observed low-temperature onset of SdH oscillations. In our samples, for the case of negligible superlattice effects, this occurs at $B \approx 0.15$ T, corresponding to an energy resolution of 0.3 meV. In lowest order the expected subband splitting in turn oscillates with magnetic field and vanishes at the flat band condition. Furthermore at the superlattice periods currently realized, which are still larger than the Fermi wavelength, screening reduces the effective superlattice amplitude substantially. In quantizing magnetic fields, screening increases in proportion to the density of states at the Fermi level and thus is expected to be further enhanced at SdH maxima [9]. Practically this means that for the superlattices considered here, in which magnetic-field-dependent screening is important, one cannot simply increase the magnetic field to enhance at a given level broadening the splitting into subbands, because that in turn reduces the effective superlattice potential and hence the splitting. These considerations demonstrate that the visibility conditions for the splitting of Landau bands into subbands are fragile and restricted to a rather narrow magnetic field regime between the onset of SdH oscillations and the regime of strong magnetic quantization. Furthermore, the superlattice unit cell must be chosen that in this magnetic field regime only a few flux quanta penetrate the unit cell, i.e. p/q is of order unity. Finally the effective potential amplitude must be tuned to a value as high as possible but below the splitting of adjacent Landau bands $\hbar\omega_c$ to avoid band overlap. With all these limitations it becomes clear that the pursuit of the Hofstadter butterfly in lateral superlattices is a tedious task.

Figure 4 shows, with two characteristic examples, that in spite of all these difficulties we have succeeded in a first observation of the splitting of Landau bands into subbands in quantitative agreement with the Hofstadter prediction. Figure 4(a) displays SdH oscillations of the magnetoresistance in a magnetic field regime in which

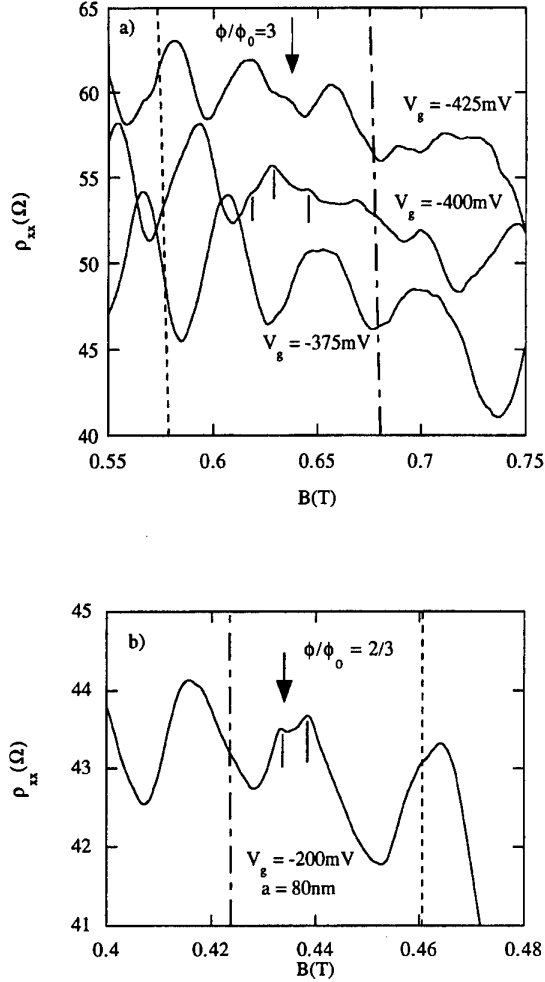


Figure 4. (a) Splitting of an SdH maximum around $\phi/\phi_0 = 3$ at $V_g = -400$ mV; at lower gate voltage (-375 mV) the potential amplitude is too small to resolve these structures whereas at higher gate voltage (-425 mV) the Landau bands begin to overlap. (b) Splitting of an SdH maximum at $\phi/\phi_0 = 2/3$ observed in an 80 nm period sample.

about three flux quanta fill the unit cell of a square superlattice with period $a = 139$ nm. The vertical broken line indicates flat band and the dot-dashed line the magnetic field where classically one expects the maximum band width. At the lowest potential modulation, $V_g = -375$ mV, one observes rather unperturbed SdH oscillations which break up into substructure at higher potential modulation. Around $\phi/\phi_0 = 3$, three distinct maxima can be identified but become less clear at higher potential modulation at which this particular Landau band further approaches the flat band case. Similarly we observe in this device splitting into p subbands at $\phi/\phi_0 = p/q = 2$ and 4 at lower and higher magnetic fields [6] respectively if we tune the superlattice amplitude accordingly to a lower or higher value than in figure 4(a). The magnetic field regime in which we expect, with the present wafer material, glimpses at the butterfly (about $0.3 \text{ T} < B < 1 \text{ T}$) corresponds

Landau subbands generated by an electrostatic superlattice

to lower flux per unit cell if we decrease the superlattice period. This allows us to observe subband splittings at fractional filling factors as exemplified for the sample with lowest period $a = 80$ nm at $\phi/\phi_0 = p/q = 2/3$ in figure 4(b). In accordance with the Hofstadter prediction we observe a Landau band to split into $p = 2$ subbands. Similarly we observe the expected splittings at $\phi/\phi_0 = 3/4$ and $5/6$, and in a 105 nm sample at $\phi/\phi_0 = 3/2$ [6].

5. Conclusion and perspectives

The good agreement with the Hofstadter prediction of flux quantum commensurability observed on several samples demonstrates beyond doubt that the observed splittings reflect the expected formation of Landau subbands and are not artifacts that could be caused, for example, by sample inhomogeneities. Nevertheless, our observations make clear that a further detailed view at the Hofstadter butterfly that may reveal the self-similar nature of the energy spectrum will be very difficult in 2DEGs on semiconductors because of the complex interplay of the superlattice effects with magnetic-field-dependent screening. Homogeneous samples with high mobilities and significantly smaller periods that are comparable to or smaller than the Fermi wavelength or the magnetic length are necessary to reveal a distinctly clearer view than the above glimpse at the Hofstadter butterfly. To create those samples remains a challenge to experimental ingenuity.

Acknowledgments

With pleasure we thank in particular F Claro, A Efros, R Ketzmerick, A Lorke, R Schuster, M Suhrke, S Ulloa and D Wharam for stimulating discussions. Financial support by the Deutsche Forschungsgemeinschaft and the ESPRIT Basic Research Action is gratefully acknowledged.

References

- [1] Hofstadter D R 1976 *Phys. Rev. B* **14** 2239
- [2] Harper P G 1955 *Proc. R. Soc. A* **68** 874
Azbel M Ya 1964 *Sov. Phys.-JETP* **19** 634
Petschel G and Geisel T 1993 *Phys. Rev. Lett.* **71** 239
- [3] Weiss D, von Klitzing K, Ploog K and Weimann G 1989 *Europhys. Lett.* **8** 179
- [4] Gerhardts R R, Weiss D and von Klitzing K 1989 *Phys. Rev. Lett.* **62** 1173
Winkler R W, Kotthaus J P and Ploog K 1989 *Phys. Rev. Lett.* **62** 1177
Beenakker C W J 1989 *Phys. Rev. Lett.* **62** 2020
- [5] Gerhardts R R, Weiss D and Wulf U 1991 *Phys. Rev. B* **43** 5192
Lorke A, Kotthaus J P and Ploog K 1991 *Phys. Rev. B* **44** 3447
- [6] Schlösser T, Ensslin K, Kotthaus J P and Holland M 1996 *Europhys. Lett.* **33** 683; *Surf. Sci.* **361/362** 847
- [7] Skuras E, Holland M, Barton C J, Davis J H and Long A R 1995 *Semicond. Sci. Technol.* **10** 992
- [8] Pfannkuche D and Gerhardts R R 1992 *Phys. Rev. B* **46** 12 606
- [9] Wulf U, Gudmundsson V and Gerhardts R R 1988 *Phys. Rev. B* **38** 4218

Low-dimensional systems in ultra-high magnetic fields: magnetic-field-induced type I to type II transitions in short-period semiconductor superlattices

N Miura[†], Y Shimamoto[†], Y Imanaka[†], H Arimoto[†], H Nojiri[†],
H Kunimatsu[‡], K Uchida[‡], T Fukuda[‡], K Yamanaka[‡],
H Momose[‡], N Mori[‡] and C Hamaguchi[‡]

[†] Institute for Solid State Physics, University of Tokyo, Roppongi, Minato-ku, Tokyo 106, Japan

[‡] Department of Electronics, Osaka University, Suita, Osaka, Japan

Abstract. We present a review on the recent study of the type I to type II transition in short-period superlattices of GaAs/AlAs by means of cyclotron resonance and interband magneto-optical spectroscopy in pulsed high magnetic fields up to 500 T. In the magneto-photoluminescence spectra of excitons in $(\text{GaAs})_m(\text{AlAs})_n$, the magnetic-field-induced type I to type II transition was observed with and without the simultaneous application of high pressure. The behaviour of the transition varies depending on the thickness of the AlAs layers. In cyclotron resonance of $(\text{GaAs})_n(\text{AlAs})_n$, the resonance peak at the X minima was observed in the type II regime for n smaller than 14, whereas the resonance at the Γ point was observed for $n > 15$. It was found that the angular dependence of the peak position does not obey the simple cosine dependence due to the subband mixing in high magnetic fields. From the angular dependence, the effective masses at the X point were determined. In high-field cyclotron resonance measurements at 129 meV up to 400 T for $n = 16$ (type I), the resonance of the X minima expected at around 260 T was indiscernible, despite the fact that the transition should have occurred at lower fields.

1. Introduction

Recently, progress in pulsed high magnetic field technology has enabled us to perform accurate measurements of the properties of matter in very high magnetic fields [1, 2]. The frontier of the high field has reached several megagauss, exceeding 500 T with the use of electromagnetic flux compression [1]. Since they make the cyclotron energy $\hbar\omega_c$ so high and the cyclotron orbital radius $(\hbar/eB)^{1/2}$ so small, the megagauss fields allow us to study various properties of nanostructure semiconductor devices with very high resolution and with a significant interplay between the magnetic field and the artificially introduced quantum potentials. They also realize the condition that $\hbar\omega_c$ exceeds various energy gaps in low-dimensional systems. Thus, various types of magnetic-field-induced level cross-over take place in very high fields, as are actually observed using the megagauss facilities in Tokyo. In GaP/AlAs short-period superlattices, for instance, a transition from the pseudo-direct gap system to the indirect systems

occurs, which was observed as a dramatic decrease of the photoluminescence intensity in high magnetic fields [3]. In InAs/GaSb short-period superlattices, a decrease of the cyclotron resonance intensity was found due to the semiconductor to semimetal transition [4]. In CdTe/CdMnTe quantum wells, a type I to type II transition was observed due to the giant Zeeman splitting in the barrier layers [5].

The magnetic-field-induced type I to type II transition in GaAs/AlAs short-period superlattices is one of such level cross-overs which can be realized at high fields. Many experimental and theoretical studies have been made on the type I to type II transition of the GaAs/AlAs system by varying the number of monolayers and applying high pressure [6–12]. It is known that the $(\text{GaAs})_n/(\text{AlAs})_n$ short-period superlattice (n is the number of monolayers) is a type I system with the conduction band minimum at the Γ point for n larger than 12–14, and the system has a direct-gap nature. On the other hand, for smaller n , it is a type II system having the conduction band minimum at the X point, with an indirect gap nature [6, 8]. The type I

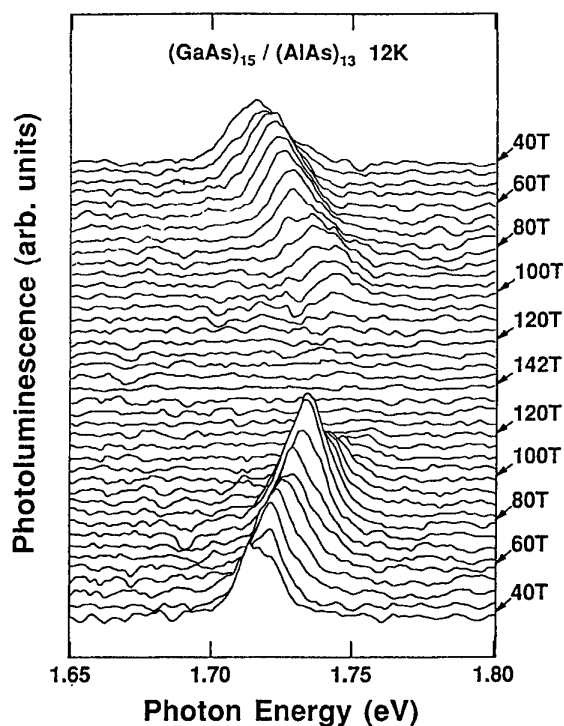


Figure 1. MPL spectra in $(\text{GaAs})_{15}/(\text{AlAs})_{13}$ up to 150 T. The data were obtained as a streak photograph in a one-shot pulse field. The time proceeds from top to bottom.

system undergoes a transition to the type II system by the application of high pressure due to the different pressure coefficients between the two minima. A similar type I to type II transition can also be induced by the application of high magnetic fields, because the effective mass at the Γ point is much smaller than those at the X point.

Such a magnetic-field-induced transition in the GaAs/AlAs system was first observed by Sasaki *et al.* in the magneto-photoluminescence (MPL) spectra [12]. Figure 1 shows the MPL spectra in $(\text{GaAs})_{15}/(\text{AlAs})_{13}$. It can be clearly seen that the MPL peak from the exciton in the Γ minimum diminishes from about 55 T with increasing field and it reappears when the field is decreased again. It was also found that there is a large hysteresis in the up and down sweeps of the field. This is due to the finite relaxation time of the electron transfer, which is comparable with the fast field sweep rate [12].

In this paper, we discuss the magnetic-field-induced type I to type II transition by means of cyclotron resonance (CR) and the MPL spectra. In the latter case, we relied upon simultaneous application of high magnetic fields and high pressures to facilitate the transition within the available field range in a wide range of samples.

2. Experimental technique

Pulsed high magnetic fields can be produced by three different techniques. First, ultra-high magnetic fields up to 550 T can be produced by electromagnetic flux compression, and secondly, we can produce a field of up to

about 200 T by the single-turn coil technique. Both of these techniques for fields exceeding 100 T are destructive and the rise time of the field to the top is a few microseconds [1]. The advantage of the latter method is that the samples and the cryostats are not destroyed by the shots of pulses, so that we can repeat the measurements on the same sample. For high-precision measurements in the lower field range, we also employ conventional non-destructive long-pulse magnets to produce fields of up to about 50 T. The rise time of the field in this case is about 10 ms.

For CR experiments, molecular gas lasers were employed as radiation sources at various wavelengths. Cooled semiconductor detectors were used for the fast detection of the transmission change. For MPL measurements, two different systems were employed depending on the magnetic field pulse duration. In millisecond long-pulse fields up to 50 T, an optical multichannel analyser (OMA) was used, and the gate of the OMA was opened at the top flat part of the field pulse to measure the spectra at constant fields. In microsecond short-pulse fields in the megagauss range, an image converter camera was employed to obtain streak pictures of the spectra, and entire magneto-spectra can be obtained in one shot of the pulse. An Ar laser was employed for the excitation at a wavelength of 514.5 nm.

The short-period superlattice samples were grown by the MBE technique [13,14]. For the CR measurements, Si was doped in the GaAs layers to provide conduction electrons of the order of 3×10^9 – $2 \times 10^{11} \text{ cm}^{-2}/\text{period}$. For the MPL measurements, non-doped samples were grown in the range $n = 15$ –19. For applying high pressures in pulsed high magnetic fields, we developed a new piston cylinder type of clamp cell made from Cu–Be [15]. In comparison to a diamond anvil cell, the advantage of this type is that the temperature rise of the sample due to the eddy-current heating of the cell in pulsed fields is almost negligible, because of the larger distance between the metal parts and the sample. Moreover, we can mount samples as large as 2 mm in diameter. A high pressure of up to about 10 kbar (1 GPa) was achievable.

3. MPL in GaAs/AlAs

$(\text{GaAs})_{17}/(\text{AlAs})_{17}$ undergoes a transition to type II at a critical pressure p_c of 0.30 GPa. Figure 2 shows the MPL spectra for the $n = 17$ sample at a pressure of $p = 0.25$ GPa. The main peak at $B = 0$ T arises from the exciton at the Γ point. It shows a persistent blue shift due to the diamagnetic shift, and the intensity shows a rapid decrease as the field is increased above 20 T. At around 20 T, the peak position starts deviating from the diamagnetic shift at the ambient pressure. From nearly the same field, the peak intensity starts decreasing. Therefore, the transition field B_c is defined at around 20 T in this sample. Above B_c , another peak appears in the lower field corresponding to the exciton at the X point. However, this X-point exciton line diminishes at a higher field. At a pressure higher than p_c , the deviation of the peak position from the ambient pressure line and the decrease of the intensity occurs as soon as we apply magnetic fields. In an

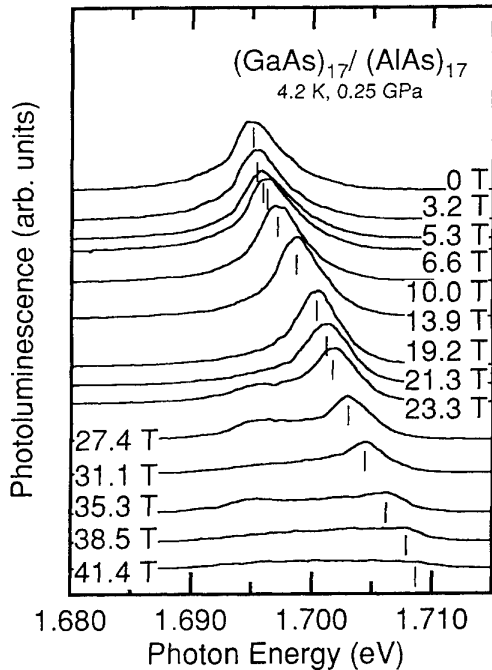


Figure 2. MPL spectra in $(\text{GaAs})_{17}/(\text{AlAs})_{17}$ under a simultaneous application of high magnetic fields and high pressure (0.29 GPa).

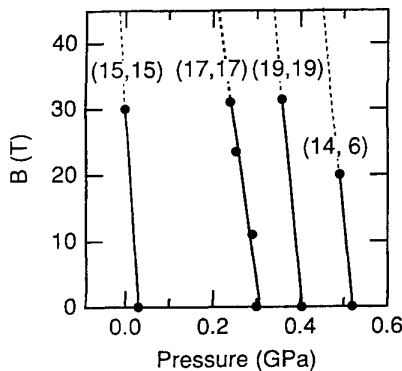


Figure 3. Phase diagram of the type I to type II transition in $(\text{GaAs})_m/(\text{AlAs})_n$ superlattices with various (m, n) on a pressure–magnetic field plane.

$n = 19$ sample whose transition pressure is $p_c = 0.40$ GPa, and an $n = 15$ sample with $p_c = 0.03$ GPa, qualitatively similar behaviour was observed. Figure 3 shows the phase diagram of the type I to type II transition point on a field–pressure plane. The gradient of the lines are almost parallel with each other indicating nearly the same effective mass of excitons and the same pressure coefficient between the samples.

The position of an exciton line at a field above B_c deviates from the type I line due to the Γ –X mixing. The intensity diminishes due to the decrease of the carrier density at the Γ point in the type II regime. These general aspects are common for $n = 15, 17$ and 19 . On the other hand, very different behaviour was observed in the

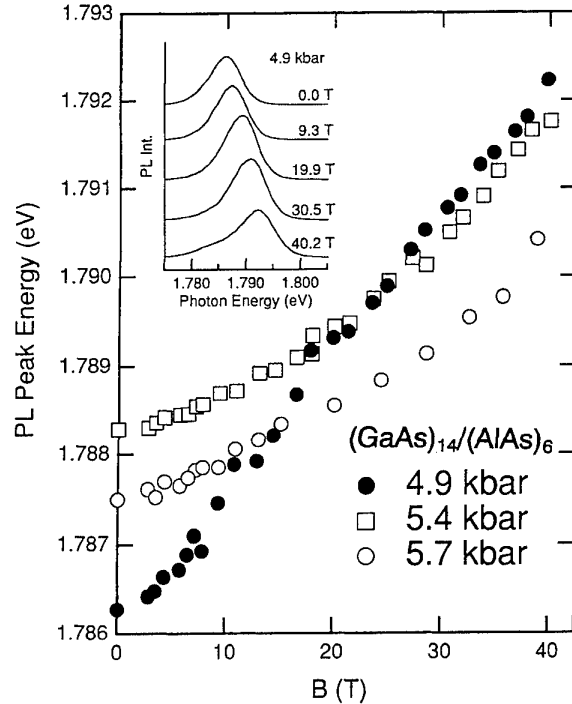


Figure 4. Diamagnetic shift of excitons in $(\text{GaAs})_{14}/(\text{AlAs})_6$ at various pressures. The inset shows the MPL spectra at $p = 0.49$ GPa. $T = 4.2$ K.

$(\text{GaAs})_{14}/(\text{AlAs})_6$ (14, 6) sample. Figure 4 shows the diamagnetic shift of the exciton below $p_c = 0.52$ GPa ($p = 0.49$ GPa), and above p_c ($p = 0.54$ GPa and 0.57 GPa). Above p_c , the exciton shows the relatively small diamagnetic shift which is characteristic of the type II exciton. At an ambient pressure, the diamagnetic shift is relatively large, corresponding to the exciton at the Γ point. At $p = 0.49$ GPa, the magnetic-field-induced type I to type II transition occurs. Namely, the diamagnetic shift changes its gradient from being Γ -like to being X-like at around 20 T. Comparing this with the $n = 15, 17$ and 19 samples mentioned above, a remarkable difference is that with increasing field, the exciton character continuously changes from Γ -like to X-like at the transition field. The intensity does not show significant change even after the transition. The difference in the behaviour between the two types of sample probably originates from the difference in the magnitude of the Γ –X mixing. In the (14, 6) sample, the mixing is large because of the thin AlAs layers, so that the X-like exciton has a large oscillator strength after the transition, whereas in samples with $n > 15$, the oscillator strength of the X-like exciton decreases as the field is increased apart from at the transition field. The intensity of the Γ -like exciton diminishes above the transition because of the small population.

4. CR in GaAs/AlAs

Cyclotron resonance was measured in doped samples with $n = 8$ – 16 in fields up to 150 T produced by the single-turn coil technique. Figure 5 shows the CR traces for these

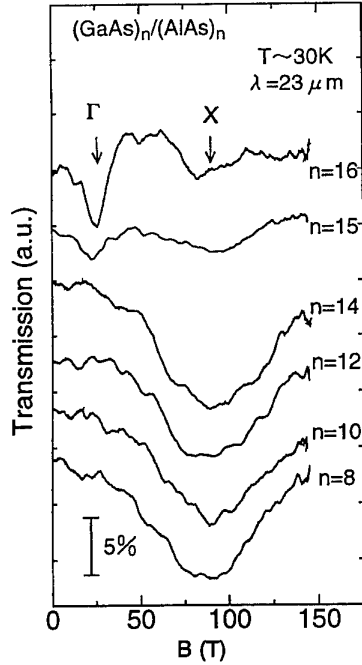


Figure 5. CR in $(\text{GaAs})_n/(\text{AlAs})_n$ superlattices.

samples at a wavelength of $23 \mu\text{m}$ ($\hbar\omega = 54 \text{ meV}$) [14]. Two resonance lines were observed. The peaks labelled Γ are associated with the Γ point and those labelled X with the X point. The Γ peak is observed for $n > 15$ and the X-peak is observed for $n < 14$. This indicates that with varying n , the type I to type II transition takes place at around $n = 14$ at $\sim 30 \text{ K}$. The effective masses for the Γ and the X peaks are $m_\Gamma = 0.053m_0$ and $m_X = 0.193m_0$ respectively. As the temperature is increased, the effective mass was found to increase. This fact implies that the transition between the donor levels has an effect at low temperatures.

When we tilted the magnetic field from the normal direction, we found that the peak from the X exciton shifts to the higher field. Figure 6 shows the angular dependence of the X-peak position of the CR [13]. It should be noted that the dependence is not cosine-like for $n = 8$, whereas it almost obeys the cosine law for $n = 14$. This indicates that for the former sample, the cyclotron energy is so large that it becomes comparable with the subband separation. In such a situation, the Landau levels are mixed with the higher subband. In fact, for $n = 14$, the interval between $N = 1$ and $N = 2$ subbands is estimated to be 44 meV , which is very close to the cyclotron energy.

The resonance peaks at low temperatures are attributed to the impurity transitions which appear at slightly lower fields than CR. We can estimate the difference between the impurity transition and the CR to be 9.4 meV . Taking account of this difference, the effective masses were obtained from the angular dependence; $m_t = 0.21m_0$ and $m_l = 1.04m_0$ for the X-point resonance. This results in $(m_t m_l)^{1/2} = 0.47m_0$, which is in very good agreement with the result of CR for bulk AlAs [16].

For $n = 16$, two peaks were observed at a wavelength of $23 \mu\text{m}$, corresponding to the Γ exciton and the X exciton.

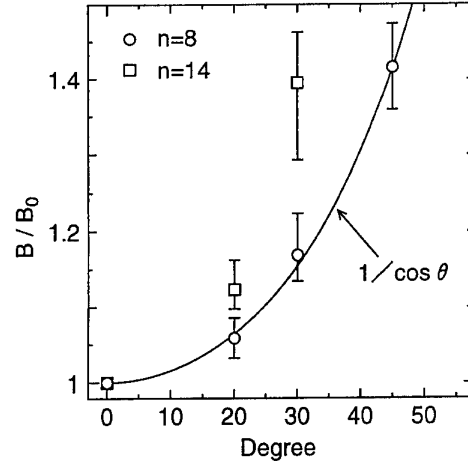


Figure 6. Angular dependence of the CR peak position at the X point in $(\text{GaAs})_8/(\text{AlAs})_8$ and $(\text{GaAs})_{14}/(\text{AlAs})_{14}$.

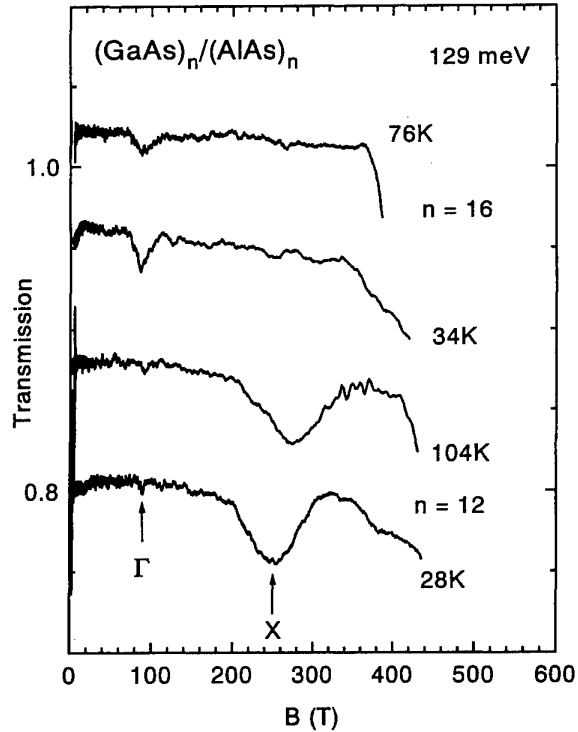


Figure 7. CR spectra in $(\text{GaAs})_n/(\text{AlAs})_n$ ($n = 12$ and 16) up to 400 T produced by electromagnetic flux compression.

This suggests that the magnetic-field-induced type I to type II transition takes place at about 60 T for $n = 16$. In order to obtain better evidence for the transition, we measured the CR at higher fields up to 400 T , using the electromagnetic flux compression. Figure 7 shows the CR spectra at 129 meV ($\lambda = 9.61 \mu\text{m}$) for $n = 12$ and $n = 16$. It should be noted that only the CR peak from the Γ point is observed for $n = 16$, whereas only the X-point peak is observed for $n = 12$. The effective masses obtained from the peak positions are $m_\Gamma^* = 0.0809m_0$ at 76 K and $m_\Gamma^* = 0.0756m_0$ at 34 K for $n = 16$, and $m_X^* = 0$.

$244m_0$ at 104 K and $m_X^* = 0.226m_0$ at 28 K for $n = 12$. According to the difference in the effective mass between the Γ point and the X point, and to the experiments up to 150 T, we should have observed the CR peak at the X point for $n = 16$ in the traces. The absence of the X-resonance peak in the data in figure 7 indicates that the type I to type II transition in the $n = 16$ sample is somehow suppressed in this experiment with very high fields.

One possible explanation for the absence of the transition is that the relaxation time of the electron transfer between the Γ and X valleys is long in comparison with the sweep of the magnetic field generated by the electromagnetic flux compression ($\sim 200 \text{ T } \mu\text{s}^{-1}$). However, it is difficult to explain the absence of the CR peak by the relaxation time of the conduction electron alone, because the cross-over effect was well observed in the case of MPL (figure 1), which ensures sufficiently fast relaxation. Another possibility is that the cross-over occurs first between the conduction band at the Γ point and the deep donor levels at the X point or the DX states [17]. These states will act as a carrier reservoir with long time constants, so that the CR at the X point may not be observed. More experimental study is necessary to clarify the participation of the DX centre, employing the subband gap light illumination.

5. Conclusions

We have investigated the magnetic-field-induced Γ -X cross-over and the type I to type II transition in $(\text{GaAs})_m(\text{AlAs})_n$ short-period superlattices by means of magneto-photoluminescence and cyclotron resonance (CR) under very high magnetic fields. In the MPL spectra, the field-induced transition was observed as a discontinuity of the diamagnetic shift and a decrease of the MPL intensity with the simultaneous application of high pressure. The characteristics of the discontinuity at the transition were found to be different depending on the barrier layer thickness.

In the CR spectra, the CR peaks corresponding to the conduction-band minima at the X point were observed in samples with the number of monolayers n smaller than 15. As the temperature was lowered, the peak position was

found to shift to lower fields due to the contribution of the impurity transition. From the angular dependence of the peak position, the effective masses at the X point were determined. In the CR spectra with a photon energy of 129 meV up to 400 T for a sample with $n = 16$, the CR peak at the X point was not discernible at an expected field around 260 T, which is far beyond the Γ -X cross-over field.

References

- [1] Miura N 1994 *Physica B* **201** 40
- [2] Miura N 1995 *Proc. 22nd Int. Conf. Phys. Semiconductors* ed D J Lockwood (Singapore: World Scientific) pp 1111
- [3] Uchida K, Miura N, Kitamura J and Kukimoto H 1996 *Phys. Rev. B* **53** 4809
- [4] Barnes D J, Nicholas R J, Warburton R J, Mason N J, Walker P J and Miura N 1994 *Phys. Rev. B* **49** 10474
- [5] Kuroda S, Kojima K, Kobayashi K, Uchida K, Miura N and Takita K, 1995 *Proc. 11th Int. Conf. High Magnetic Fields in Semiconductor Physics (Cambridge, 1994)* ed D Heiman (Singapore: World Scientific) p 662
- [6] Cingolani R, Holtz M, Muralidharan R, Ploog K, Reimann K and Syassen K 1990 *Surf. Sci.* **228** 217
- [7] Li G, Jian D, Han H, Wang Z and Ploog K 1990 *J. Lumin.* **46** 261
- [8] Matsuoka T, Nakazawa T, Ohya T, Tcniguchi K, Hamaguchi C, Kato H and Watanabe Y 1991 *Phys. Rev. B* **43** 11798
- [9] Ihm J 1987 *Appl. Phys. Lett.* **50** 1068
- [10] Nakayama M, Tanaka I, Kimura I and Nishimura H 1990 *J. Phys. Soc. Japan* **29** 41
- [11] Xia J B and Chang Y C 1990 *Phys. Rev. B* **42** 1781
- [12] Sasaki S, Miura N, Yagi T and Horikoshi Y 1993 *J. Phys. Soc. Japan* **62** 2490
- [13] Yamanaka K, Momose H, Mori N, Hamaguchi C, Arimoto H, Imanaka Y, Shimamoto Y and Miura N 1996 *Physica B* at press
- [14] Fukuda T, Yamanaka K, Momose H, Mori N, Hamaguchi C, Imanaka Y, Shimamoto Y and Miura N 1995 *Proc. EP2DS (Nottingham, 1995)*
- [15] Miura N, Kunimatsu H, Uchida K, Sasaki S and Yagi T 1995 *Proc. 11th Int. Conf. High Magnetic Fields in Semiconductor Physics (Cambridge, 1994)* ed D Heiman (Singapore: World Scientific) p 324
- [16] Miura N, Yokoi H, Kono J and Sasaki S 1991 *Solid State Commun.* **79** 1039
- [17] Wilamowski Z, Kossut J, Jantsch W and Ostermayer G 1991 *Semicond. Sci. Technol.* **6** B38

Intersubband dynamics of asymmetric quantum wells studied by THz ‘optical rectification’

K Unterrainer[†], J N Heyman[‡], K Craig[§], B Galdrikian[§],
M S Sherwin[§], K Campman[§], P F Hopkins[§] and A C Gossard[§]

[†] Institute for Solid State Electronics, Technical University Vienna, A-1040 Vienna, Austria

[‡] Department of Physics, Macalester College, Saint Paul, MN 5505, USA

[§] University of California at Santa Barbara, Santa Barbara, CA 93106, USA

Abstract. We have measured the rectification of far-infrared radiation resonant with the lowest intersubband transition of an AlGaAs/GaAs asymmetric coupled double-quantum well in which the subband spacing is 11 meV. From these measurements we can extract an intersubband lifetime of 1.2 ± 0.4 ns at low excitation intensity and $T = 10$ K, which appears promising for devices such as FIR detectors or mixers which can operate at low excitation and temperature. In order to investigate the effect of carrier concentration on the relaxation time we have performed the same experiments in a logarithmically graded quantum well.

1. Introduction

The recent development of a mid-infrared intersubband laser has encouraged the study of the possibility of an intersubband laser in the far-infrared spectral region which has a great demand for semiconductor radiation sources [1]. The most important parameter for such devices, as well as for detectors, is the intersubband relaxation time (T_1). We have measured this time using a new experimental technique based on the strong nonlinear susceptibility of an asymmetric double quantum well. Simultaneous measurements of the rectification of far-infrared radiation resonant with the intersubband transition and of the intersubband absorption enable the determination of the intensity-dependent relaxation time. Optical rectification is the static polarization produced by difference-frequency mixing of a harmonic electric field $E(\omega)$ with itself through the second-order susceptibility. In asymmetric quantum wells this static polarization is observed for intersubband transitions, because the expected value of the position of an electron in an excited state and in the ground state are different

$$P^{(2)}(0) = \chi_{(0;\omega,\omega)}^{(2)} E(\omega) E(\omega) = -e(n_2 - n_2^0)(z_{2,2} - z_{1,1}) \quad (1)$$

where n_2 is the population in the second subband, n_2^0 is the equilibrium population in the second subband and $z_{i,j}$ is the dipole-matrix element between states i and j .

2. Experiment

The coupled GaAs quantum wells in our heterostructure are 85 Å and 73 Å wide, separated by a $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ barrier 25 Å wide (see figure 1(a)). The expectation value for the electrons in the ground state is larger in the wide well and that for the electrons in the first excited state is larger in the narrow well (see figure 1, 0 V gate bias). The lowest electron-subband spacing was measured to be $E_2 - E_1 = 11$ meV using photoluminescence. Energies of the higher subbands $E_3 - E_1 = 110$ meV and $E_4 - E_1 = 156$ meV were obtained from a self-consistent model described below. The measurements described below were performed near resonance with $h\nu \approx E_2 - E_1$. Under these conditions the heterostructure may be approximated as a two-subband system. At low temperatures ($T < 50$ K), the zero-bias charge density obtained by Hall measurements is $N_s = 2 \times 10^{11} \text{ cm}^{-2}$ and the electron in-plane mobility is $\mu = 1 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. In this paper we have performed a comparative study of the relaxation in the double quantum well and in a logarithmically graded quantum well. The wide graded logarithmic well was designed with the specific goal of a widely tunable FIR response. The 1455 Å logarithmic well was constructed from a 15 Å period GaAs/AlGaAs digital alloy superlattice, with the average Al composition x per period varied so that the logarithmic shape was achieved. Use of a surface gate, an epitaxially grown backgate, and contacts to the electron gas in the quantum well allow independent control of the absorption frequency and of the carrier concentration N_s [2].

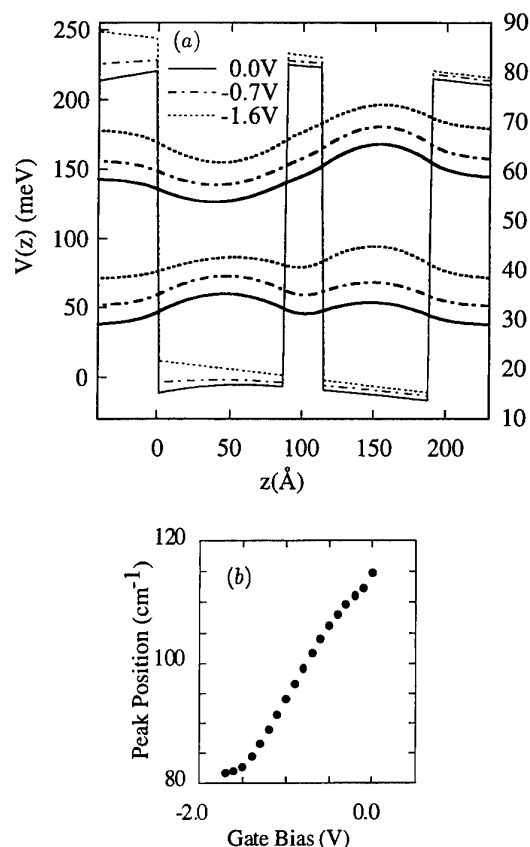


Figure 1. (a) Conduction band diagram for the coupled double well for different gate bias voltages. The envelope wavefunctions for the first and second subband are shown (their vertical positions are not drawn to scale). (b) Intersubband absorption frequency as a function of gate bias.

The absorption spectrum of the double well consists of a single line with an approximately Lorentzian lineshape, and the peak position can be tuned between 14.5 and 10 meV (115 cm^{-1} – 80 cm^{-1}) with gate bias (see figure 1(b)). At 0 V bias the peak position is $\hbar\omega_{1,2} = 14.3\text{ meV}$ (115 cm^{-1}), and the full-width at half maximum is 0.55 meV. The difference between the energy of the absorption resonance and the subband spacing is due to the depolarization shift. Resonant harmonic generation measurements on this structure have been reported elsewhere [3].

Our optical rectification experiments employ the UCSB Free-Electron Laser (FEL) as a far-infrared pump. The laser was tuned to produce $5\text{ }\mu\text{s}$ pulses of 103 cm^{-1} radiation at a 1.5 Hz repetition rate. The FIR uniformly illuminates the edge of our sample, which is mounted in a variable-temperature cryostat. Optical rectification is observed as a laser-induced voltage between the gate and the quantum well, because the expected value of the position of an electron in an excited state and the ground state are different. The impedance of the detection circuit and of the gate bias voltage source was high to ensure that the time constant was longer than the pulse width. Radiation transmitted through the sample is focused onto a 4.2 K bolometer, so that intersubband absorption and

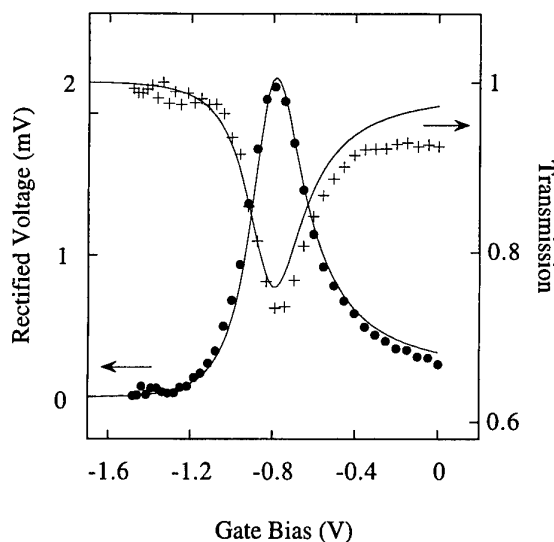


Figure 2. Optical rectification and transmission in the linear regime ($T = 10\text{ K}$, $I = 0.5\text{ W cm}^{-2}$). Full curves are calculated from theory (see text).

optical rectification can be measured simultaneously. We investigated resonant effects by tuning the intersubband absorption resonance through the laser frequency with gate bias. Optical rectification in step quantum wells in the mid-infrared has been suggested and recently observed by Rosencher *et al* [4].

3. Results and discussion

Figure 2 shows the amplitude of the optical rectification and the FIR transmission of the double well as a function of gate bias, measured at $T = 10\text{ K}$ and at a laser intensity of 0.5 W cm^{-2} . A resonance in the rectification occurs when the intersubband absorption frequency is tuned through the laser frequency at $V_G = -0.8\text{ V}$, indicating that the rectification is associated with the intersubband transition.

We have calculated the energies and envelope functions of electrons at the subband minima in our double well by self-consistently solving the Schrödinger and Poisson equations within the effective-mass approximation. Temperature effects were included by allowing thermal population of higher subbands. This model, together with the dephasing time determined from the linewidth of the absorption resonance, allows us to calculate the FIR transmission and optical rectification of our sample (figure 2, full curves), and we find a good agreement with the experiment.

Figure 3(a) shows the sensitivity (rectified voltage/intensity) and the intersubband absorption at resonance as a function of pump intensity at $T = 10\text{ K}$. At low intensities the sensitivity and the absorption are constant. The rectified signal starts to saturate (the sensitivity decreases) at about 1 W cm^{-2} and the absorption coefficient starts to decrease at about 200 W cm^{-2} . The saturation behaviour of both the rectified signal and the absorption is described

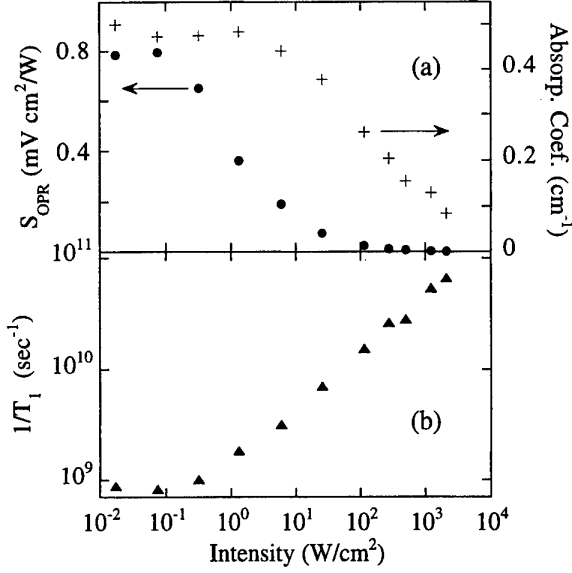


Figure 3. (a) S_{OPR} (rectified voltage divided by intensity) and intersubband absorption coefficient versus pump intensity, measured at $T = 10$ K. (b) Intersubband scattering rate calculated from these data.

by the rate equations of a two-level system. The intensity-dependent rectified signal is then given by

$$\Delta V^{\text{opr}} = \frac{-4\pi e}{\epsilon} (z_{2,2} - z_{1,1}) \Delta n_{1,2}^0 \frac{I \sigma T_1}{\hbar \omega} \left(1 + 2 \frac{I \sigma T_1}{\hbar \omega} \right)^{-1} \quad (2)$$

and the intensity-dependent absorption coefficient is

$$\alpha(I) = \Delta n_{1,2}^0 \frac{\sigma}{a} \left(1 + 2 \frac{I \sigma T_1}{\hbar \omega} \right)^{-1} \quad (3)$$

where σ is the linear absorption cross section, $\Delta n_{1,2}^0$ is the difference between the equilibrium populations of the first and second subbands, I is the intensity of the pump beam and a is the thickness of the sample. Fitting the magnitude of the rectification voltage (equation (2)) to the experiment yields an intersubband lifetime $T_1 = 1.2 \pm 0.4$ ns. This simple fit, however, does not explain the intensity dependence for intermediate power levels very well, which indicates that the assumption of a constant relaxation time is not suitable for all intensities. The obtained lifetime is longer than the calculated lifetime of $T_1 = 215$ ps we obtain from Ferreira and Bastard's simple model for the acoustic-phonon scattering rate in our structure [5]. In a wide modulation doped square well Craig *et al* have also found relaxation times in the nanosecond range [6]. Since the intersubband excitation is in fact a collective excitation, the relaxation of a collective mode has to be treated in a different way from the single-particle relaxation. This many-body effect can reduce the relaxation rate [7].

The different saturation behaviour of the rectified signal and of the absorption coefficient indicates that the relaxation time is intensity dependent. This intensity dependence has not been considered in any other experiment. Moreover,

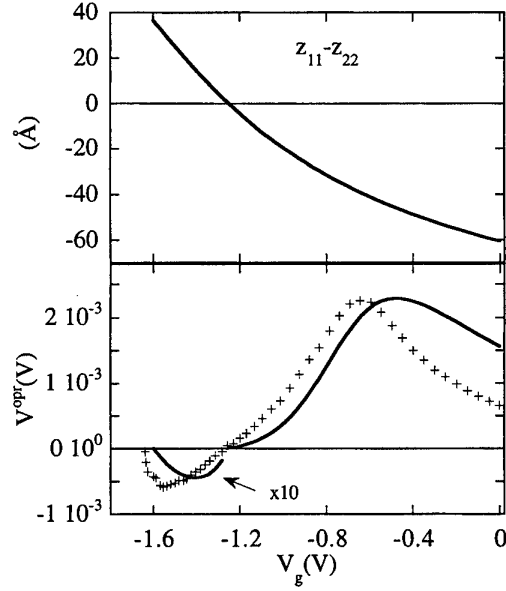


Figure 4. Upper part: $z_{2,2} - z_{1,1}$ as a function of gate bias voltage as obtained from a self-consistent calculation. Lower part: crosses show the measured rectified voltage versus gate voltage for an intensity of $I = 45$ W cm^{-2} at $T = 10$ K. The full curve shows a calculation using the parameters of our self-consistent simulation and $T_1 = 1.2$ ns.

most experimental results are derived from nearly saturated systems. However, the intensity-dependent relaxation time can be obtained directly from the ratio between rectified signal and the absorption coefficient (see equations (2) and (3)) [8]

$$T_1 = - \frac{V^{\text{opr}}(I)}{I \alpha(I)} \frac{\epsilon \hbar \omega}{4\pi e (z_{2,2} - z_{1,1}) a} \quad (4)$$

Therefore, simultaneous measurements of optical rectification and intersubband absorption can determine T_1 over a wide range of conditions, including low intensities within the linear regime. The intersubband lifetime depends strongly on intensity (figure 3(b)). In the linear regime we measure $T_1 = 1.2$ ns (consistent with the result of the fit only to the rectified signal), while in the strongly saturated regime ($I = 2$ kW cm^{-2}) we find $T_1 = 15$ ps. This intensity dependence helps to explain the large scatter of experimentally obtained values for the relaxation time below the optical phonon interaction. Most experiments were done at high intensities resulting in short T_1 s, except that by Faist *et al* which also reports quite long relaxation times [9]. Although pump and probe experiments have the advantage of direct determination of T_1 without the need for an absolute intensity calibration, their drawback is that these experiments have to be performed at high intensities with a significant saturation. The low-power limit is not accessible in pump and probe experiments. Thus, Murdin *et al* observed a relaxation time of about 50 ps in a homogeneously 5×10^{15} cm^{-3} doped multi-quantum well [10]. This value has to be compared with our results at intermediate and high pump intensities. We suggest two mechanisms for the reduction in lifetime at high intensities.

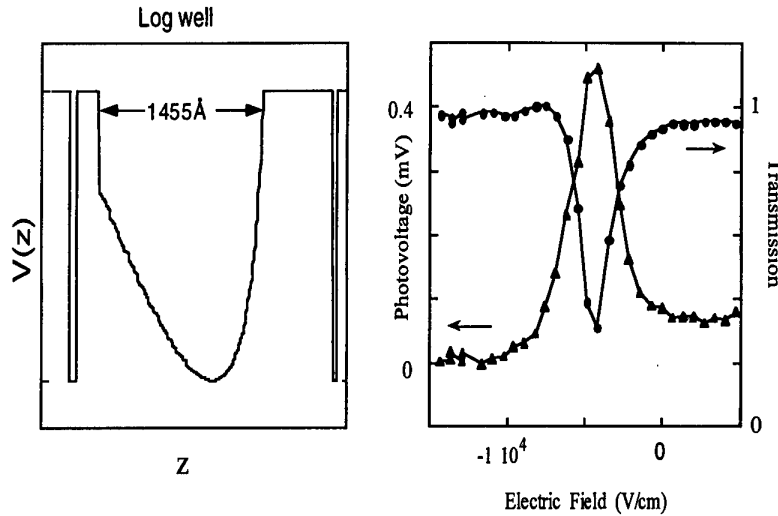


Figure 5. Left: band diagram of the logarithmically graded quantum well. Right: optical rectification and transmission in the linear regime ($T = 10$ K, $I = 0.5$ W cm $^{-2}$).

First, the electron–electron intersubband-scattering rate is predicted to depend strongly on the population in the second subband. If this mechanism controls the relaxation, the lifetime will decrease as the excited state is populated through absorption. (At highest intensities the population in the first excited subband reaches 50% of the total electron concentration.) Second, the excitation by the pump should heat the electron gas. In view of the strong temperature dependence of the scattering rate, a reduction in the intersubband lifetime at high pump intensities is expected due to heating.

Figure 4 shows in the lower part the rectified signal as function of applied bias at an intensity ($I = 45$ W cm $^{-2}$) which is higher than the resonant saturation intensity ($I > I_s$) assuming $T_1 = 1.2$ ns. The power broadening of the rectified signal is apparent from a comparison with figure 2. At a gate bias of $V_g = -1.3$ V a sign change of the rectified signal occurs. This sign change of the rectified signal can only be explained by a sign change of $z_{2,2} - z_{1,1}$. The upper part of figure 4 shows the dependence of $z_{2,2} - z_{1,1}$ obtained from our simulations. The calculated sign change of $z_{2,2} - z_{1,1}$ and of the measured rectified signal correspond very well. The reason for this sign change is that the centre of the wavefunctions switch between the wells at $V_g = -1.3$ V. This is also evident from figure 1 where the wavefunctions are shown. For $V_g = 0$ V, and $V_g = -0.68$ V the centre of the wavefunction of the lowest subband is in the wide well. At $V_g = -1.6$ V the centre of this wavefunction is centred in the narrow well. The symmetry of the wavefunction of the second subband is not influenced by gate bias.

In addition, we have used equation (5) to calculate the rectified signal at this intensity as a function of gate bias (see the full curve in the lower part of figure 4). The overall agreement with the experimental data is good. There are slight discrepancies for the lineshape of the resonance. We think these are caused by an intensity-dependent relaxation

time, and a theoretical value for $z_{2,2} - z_{1,1}$ at low negative bias voltage which is too large.

In order to investigate the effect of carrier concentration N_s on the relaxation time we performed the same experiments in a logarithmically graded quantum well. Using the advanced backgate technique of the logarithmic well we were able to change the carrier concentration while keeping the absorption frequency constant. The results from the logarithmic well show a considerable strong rectified signal (figure 5). The resonance frequency is tunable between 35 cm $^{-1}$ and 125 cm $^{-1}$ by the gate voltage, demonstrating that this structure operates as widely voltage tunable FIR ‘photo’-detector.

The measurements at different carrier concentrations at fixed absorption frequency show that the integrated absorption increases, which is expected (figure 6). However, the rectified signal decreases with increasing carrier concentration. The evaluated corresponding relaxation times are shorter than that of the double well. This difference can be caused by alloy scattering in the logarithmic well. The dependence on carrier concentration shows a decrease with increasing carrier concentration from 300 ps at $N_s = 5 \times 10^{10}$ cm $^{-2}$ to 25 ps at $N_s = 2 \times 10^{11}$ cm $^{-2}$. This behaviour is also true at very low intensities where we expect no heating of the electron gas. Therefore, we can conclude that the electron–electron scattering increases the relaxation, and the assumption that the relaxation is slowed down by the collective mode is not supported. At higher intensities we cannot separate the influence of electron–electron scattering from a heating effect of the electron gas since the increased absorption at higher carrier concentrations causes more power to be absorbed giving rise to stronger heating.

In summary, we have observed resonant optical rectification in our coupled quantum well. From simultaneous measurements of the rectified voltage and of the absorption we can extract the intersubband lifetime as a function of intensity. At low intensities the lifetime is $1.2 \pm$

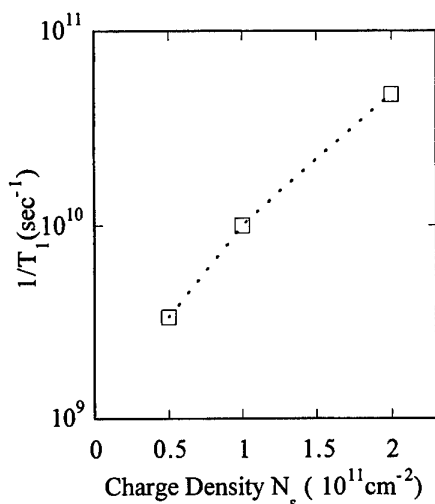


Figure 6. Charge density dependence of the relaxation rate in a logarithmically graded quantum well.

0.4 ns at $T = 10$ K which appears promising for devices which can operate at low excitation and temperature, such as FIR detectors or mixers. At high intensities the lifetime reduces to 15 ps at 2 kW cm^{-2} which is still an order of magnitude larger than the lifetime for intersubband energies above the optical phonon energy. The results of the logarithmic well show that, based on optical rectification, a widely electric field tunable FIR detector can be realized. In addition, we have shown that the electron-electron scattering increases the intersubband relaxation time.

Acknowledgments

The authors would like to thank the staff at the Center for Free-Electron Laser Studies, J R Allen, D Enyeart, J P Kaminsky, G Ramian and D White. Funding for the Center for Free-Electron Laser studies is provided by the Office of Naval Research.

References

- [1] Faist J, Capasso F, Sivco D L, Sitori C, Hutchinson A L and Cho A Y 1994 *Science* **264** 553-6
- [2] Hopkins P F, Campman K L, Bellomi G, Gossard A C, Sundaram M, Yuh E L and Gwinn E G 1994 *Appl. Phys. Lett.* **64** 348
- [3] Heyman J N, Craig K, Galdrikian B, Campman K, Hopkins P F, Fafard S, Gossard A C and Sherwin M S 1994 *Phys. Rev. Lett.* **72** 2183-6
- [4] Rosencher E, Bois P, Costard E, Delaitre S and Nagle J 1989 *Appl. Phys. Lett.* **55** 1597
- [5] Ferreira R and Bastard G 1989 *Phys. Rev. B* **40** 1074
- [6] Craig K, Galdrikian B, Heyman J N, Markelz A G, Sherwin M S, Campman K, Hopkins P F and Gossard A C *Phys. Rev. Lett.* at press
- [7] See, for example, Nozieres P and De Dominicis C T 1969 *Phys. Rev.* **178** 1097
- [8] Heyman J N, Unterrainer K, Craig K, Galdrikian B, Sherwin M S, Campman K, Hopkins P F and Gossard A C 1995 *Phys. Rev. Lett.* **74** 2682
- [9] Faist J, Sitori C, Capasso F, Pfeiffer L and West K 1994 *Appl. Phys. Lett.* **64** 872
- [10] Mordin B N, Knippels G M H, van der Meer A F G, Pidgeon C R, Langerak C J G M, Helm M, Heiss W, Unterrainer K, Gornik E, Geerink K K, Hovenier N J and Wenckebach W T 1994 *Semicond. Sci. Technol.* **9** 1554

Virtual states, dynamic localization, absolute negative conductance and stimulated multiphoton emission in semiconductor superlattices

B J Keay[†], C Aversa[†], S Zeuner[†], S J Allen Jr[†], K L Campman[‡],
K D Maranowski[‡], A C Gossard[‡], U Bhattacharya[§] and
M J W Rodwell[§]

[†] Center for Free-electron Laser Studies, UCSB, Santa Barbara, CA 93106, USA

[‡] Materials Department, UCSB, Santa Barbara, CA 93106, USA

[§] Department of Electrical and Computer Engineering, UCSB, Santa Barbara, CA 93106, USA

Abstract. We report the observation of dynamic localization, absolute negative conductance (ANC) and multiphoton stimulated emission in sequential resonant tunnelling semiconductor superlattice bow-tie antenna coupled to intense terahertz electric fields. Perhaps the most remarkable observation is that with increasing terahertz field strength the conductance near zero d.c. bias decreases towards zero and then becomes negative. The results presented here compare favourably with a model in which virtual states, familiar from nonlinear optics, take a role similar to the unperturbed quantum well states.

In this paper we report the observation of dynamic localization, absolute negative conductance (ANC) [1,2] and multiphoton stimulated emission in multi-quantum well semiconductor superlattices. We believe that the observations reported here are due to photon-assisted tunnelling (PAT) between the ground states of neighbouring quantum wells. This result is particularly surprising because it was believed that a ladder of levels, the ground states of superlattice quantum wells in an applied electric field, are all equally populated and therefore offer no net coupling to the photon field [3].

Photon-assisted tunnelling (PAT), a well established phenomenon describing quasi-particle tunnelling in superconducting electronics, was first described theoretically in 1963 by Tien and Gordon [4]. They showed that a solution to Schrödinger's equation could be found in the case of an electron in a superconducting tunnelling junction coupled to an a.c. electric field. More specifically, if the perturbation, $eV \cos \omega t$, is defined to be the applied potential difference between two superconducting films, then the relevant solution to Schrödinger's equation was shown to be

$$|a(t)\rangle = |a\rangle \sum_{n=-\infty}^{\infty} \{J_n(\alpha) \exp[-i(E_a + n\hbar\omega)t/\hbar]\} \quad (1)$$

where V is a constant, $\alpha = eV/\hbar\omega$ and $J_n(\alpha)$ are Bessel functions. Here, for purposes of discussion, we will consider two quantum well states. The state $|a\rangle$ we will

assume to be primarily localized in the left well with an unperturbed energy E_a and the state $|b\rangle$ we will assume to be localized in the right well with unperturbed energy E_b (see figure 1). From their model Tien and Gordon (TG) were able to explain qualitative features of photon-assisted tunnelling (PAT) currents observed in superconducting diodes [4–6]. The TG model, in some sense, is the 'standard model' of PAT and has been used extensively to describe PAT in other systems [7–23]. Experimentally, PAT has also emerged recently in transport in semiconductor superlattices [24–27], multiple quantum wells [28, 29], and nanostructures [30–32] in the presence of high-frequency fields. Some of the more recent observations include multiphoton-assisted tunnelling and PAT-induced electric field domains [27]. Some of these results [6, 25, 27] lend support to the main prediction of their model, namely the Bessel function dependence of the PAT channels.

The model of Tien and Gordon is also interesting because it assumes that virtual states [33–39], which are usually neglected in standard perturbative approaches, are the dominant contribution to the PAT current. By a virtual state we mean a state that has the same spatial wavefunction, $\langle r|a\rangle$, as the unperturbed state, $|a\rangle$, but has associated energy that differs from the unperturbed energy, E_a , by some amount ΔE . Here, we are interested in the special case that the virtual states are situated at energies given by $E_a + n\hbar\omega$.

From perturbation theory, we can write the perturbed

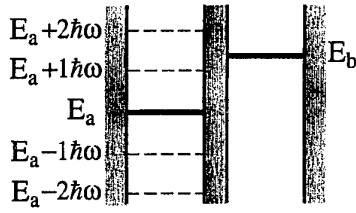


Figure 1. The two quantum wells separated by a barrier.

state as [39]

$$|a(t)\rangle = \sum_n \{ (J_n(\alpha) + A_n) \exp[-i(E_a + n\hbar\omega)t/\hbar] |a\rangle + B_n \exp[-i(E_a + n\hbar\omega)t/\hbar] |b\rangle \} \quad (2)$$

where now $\alpha = \langle a|eV|a\rangle/\hbar\omega$, and V is no longer restricted to being a constant. It is important to note that, in addition to the Bessel function contribution of Tien and Gordon [4] to the perturbed state, we have two additional coefficients, A_n and B_n . These coefficients contain terms that include at least one matrix element of the form $\langle b|eV|a\rangle$ and/or $\langle b|eV|b\rangle$. Furthermore, it can be seen that the Bessel functions contain no matrix elements connecting the state $|a\rangle$ to the state $|b\rangle$; they only contain the transition matrix element $\langle a|eV|a\rangle$. Thus, equation (1) corresponds to a quantum mechanical sum over virtual states, and the approach of Tien and Gordon is equivalent to neglecting the coupling between the quantum wells; then after obtaining expressions for the perturbed states, the coupling between the wells is introduced as an a.c. field-independent tunnelling probability, T_{ab} . The n -photon-assisted probability is then given by $J_n^2(\alpha)T_{ab}$, which we will use later in describing the observation of dynamic localization and ANC.

At this point the reader might be alarmed by the fact that it is the coefficient B_n , that, at lowest order in perturbation theory, yields Fermi's golden rule for the photon-assisted transition between $|a\rangle$ and $|b\rangle$. Thus, theories that retain only the Bessel functions neglect a class of transitions which are not restored by introducing the tunnelling probability, T_{ab} . Nevertheless, the approach of Tien and Gordon should be a reasonable approximation provided that $J_n^2(\alpha)T_{ab} \gg |B_n|^2$.

Furthermore, it should be pointed out that there is some freedom in choosing how the Bessel functions enter into the formulation discussed above. In particular, since the energy origin can always be chosen arbitrarily, a time-dependent phase (gauge) transformation allows one to choose either $\langle a|eV|a\rangle$ or $\langle b|eV|b\rangle$ to be equal to an arbitrary value; in particular, at least one of these can be chosen to be zero. The argument of the Bessel functions changes as $\langle a|eV|a\rangle$ and $\langle b|eV|b\rangle$ change under such transformations, but in the end all calculated quantities should be insensitive to these differences in formulation. Thus, we see that Tien and Gordon have made an implicit selection of the gauge such that $\langle a|eV|a\rangle$ is the difference in potential energy between the well that the electron is tunnelling from and the well that the electron is tunnelling to. This choice is equivalent to setting $\langle b|eV|b\rangle$ equal to zero. We

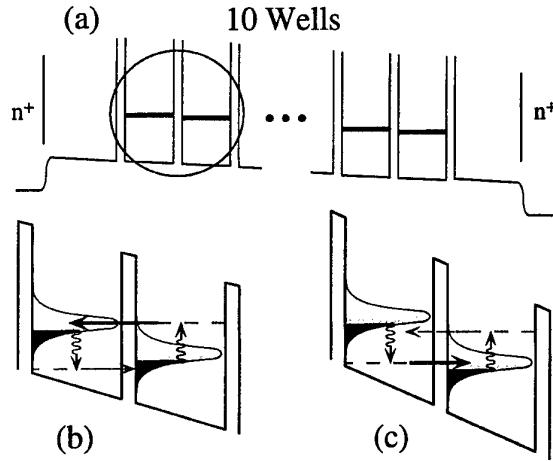


Figure 2. (a) The superlattice structure under d.c. bias. (b) Broadened levels in the presence of radiation at low d.c. bias. For $edE_{DC} < \hbar\omega$, emission channels are inhibited while the absorption channels are enhanced. (c) At higher bias, $edE_{DC} > \hbar\omega$, the absorption channels are inhibited while the emission channels are not. The vertical arrows represent fixed photon energy and the horizontal arrows indicate the resonantly enhanced process (thick arrow) and the impeded process (thin arrow) respectively.

stress that for a state $|b\rangle$ localized within the same well as $|a\rangle$, or for next-nearest neighbour transitions, this gauge dependence requires a reinterpretation of α in order to apply the Tien and Gordon model. For example, for *next*-nearest neighbour photon-assisted tunnelling we would expect to observe Bessel functions with the argument 2α instead of α .

It is not obvious that the transitions that contribute to the Bessel functions (i.e. the $\langle a|eV|a\rangle$ matrix elements), and that are important at intense laser fields, should be of fundamentally different character from those that are important at weak fields, usually attributed to the $\langle b|eV|a\rangle$ matrix elements. Nevertheless, the observations we will present here do lend support to the predicted Bessel function dependence of the PAT currents.

The sequential resonant tunnelling superlattice used in these experiments is shown in figure 2. It consisted of 300 nm of GaAs doped at $n^+ = 2 \times 10^{18} \text{ cm}^{-3}$, followed by a 50 nm GaAs spacer layer and ten 15 nm GaAs quantum wells and eleven 5 nm GaAs/ $\text{Al}_{0.30}\text{Ga}_{0.70}\text{As}$ barriers and capped by another 50 nm GaAs spacer layer and 300 nm of $n^+ = 2 \times 10^{18} \text{ cm}^{-3}$ Si-doped GaAs. The substrate is semi-insulating and the superlattice and spacer layers were n doped to $3 \times 10^{15} \text{ cm}^{-3}$ with Si. These experiments were facilitated by integrating the superlattice into a broad band bow-tie antenna. The devices were then glued onto a high-resistivity hemispherical silicon lens and gold wires were bonded to the two gold bows. The experiments were performed over a temperature range of 8–15 K with the sample mounted in a temperature-controlled flow-type cryostat with Z-cut quartz windows. Radiation was incident on the curved surface of the silicon lens with the polarization parallel to the axis connecting the two gold lobes of the bow-tie.

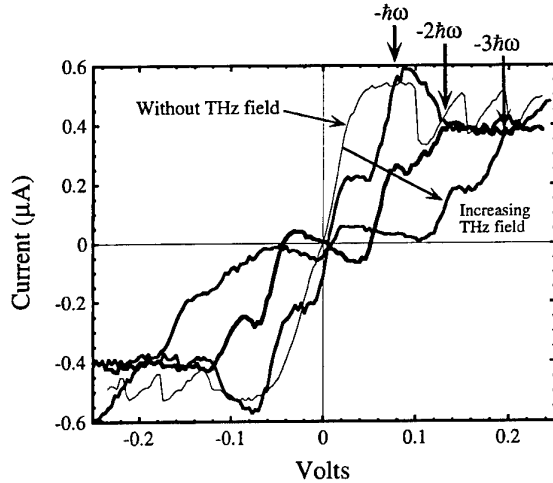


Figure 3. Current-voltage characteristics measured without radiation and with 1.30 THz (5.38 meV) radiation at three different laser field strengths.

In figure 3 we show the d.c. current-voltage (I - V) characteristics measured without radiation and with 1.30 THz radiation at three different laser powers. The I - V without laser radiation ($E_{AC} = 0$) shows an ohmic region characteristic of sequential resonant tunnelling followed by saw-tooth oscillations associated with electric field domains [27, 40]. At low bias, the current through the sample occurs via ground state to ground state tunnelling. As the bias is increased, the current approaches a 'critical current', the maximum current that the ground state to ground state tunnelling will support, and a quantum well breaks off forming a high-field domain. The high-field domain is characterized by the alignment of the ground state in one well with the excited state in the 'down-hill' well. As the bias is increased still further, one well after another breaks off into the high-field domain, indicated by the saw-tooth NDC structure [27, 40], until the entire sample is encompassed by the high-field domain. In this work we are concerned only with the behaviour near zero bias, where the transport occurs via ground state to ground state tunnelling.

Perhaps the most remarkable result is found near zero d.c. bias. In the presence of a terahertz electric field the conductance at zero bias is first suppressed as the electrons are distributed over the virtual states, a manifestation of dynamic localization. However, as the d.c. bias is increased the one-photon emission channel is brought into resonance with the ground state in the neighbouring well and the current increases, resulting in the formation of a step in the I - V characteristic. (Note that the voltage where this occurs is approximately $N\hbar\omega/e$, where N is the number of quantum wells; $N = 10$ here.)

At the next higher power shown in figure 3, the conductance near the origin actually becomes negative. When this happens the electrons use the absorbed energy from the laser field to tunnel against the applied d.c. bias. While localization and near zero conductance is predicted, absolute negative conductance is not.

At this level of terahertz excitation, as the d.c. bias voltage is increased beyond the absolute negative conductance region, the one-photon emission step is encountered in the I - V characteristic (figure 3). The current becomes positive and the electrons are driven down the superlattice accompanied by the stimulated emission of a photon [41]. Remarkably, a two-photon emission step is also seen at the appropriate voltage. Finally, at the highest power shown in figure 3, the zero-photon channel has reappeared, the one-photon emission step is substantially quenched, the current at the two-photon step has decreased by a factor of two, and the formation of the three-photon emission step is clearly seen at the appropriate bias voltage. The location in voltage of the one-photon and two-photon emission steps increases with increasing frequency, as would be expected from multiphoton stimulated emission.

We can project these results on the model of Tien and Gordon in the following way. We assume that the basic transport mechanism is the sequential tunnelling from one quantum well to its neighbour. Following Tien and Gordon [4], we expect that in the presence of a laser field the n -photon tunnelling probability is proportional to $J_n^2(\alpha)$, where n corresponds to the net 'emission' (negative n) or net 'absorption' (positive n), of n -photons, and the argument is given by $\alpha = edE_{AC}/\hbar\omega$. Here e is the electron charge, d is the superlattice period, E_{AC} is the THz field strength and $\hbar\omega$ is the photon energy. In particular, $J_0^2(\alpha)$ describes the suppression of the conductance at zero bias, i.e. dynamic localization, and we expect the conductance to be driven to zero when the argument, $edE_{AC}/\hbar\omega$, equals the zero of the $J_0(\alpha)$ Bessel function. Likewise, we expect the new channels at $n\hbar\omega$, that appear as steps at the appropriate bias voltage, to develop in a non-monotonic way as $J_n^2(\alpha)$ [27].

While the dynamic localization is expected, the absolute negative conductivity does not arise in a straightforward way from this model. As a possible explanation for these observations we propose that the symmetry between the emission and absorption processes is broken by scattering at low temperatures. This effect is illustrated in figures 2(b) and (c) where a broadened level in one well is shown, coupled by the laser field to the neighbouring well separated by edE_{DC} .

It is important to realize that near the first zero of $J_0^2(\alpha)$ transport will be dominated by the photon-assisted channels. In figure 2(b) the applied d.c. field is less than $\hbar\omega/ed$, and the one-photon emission channel is impeded, but the absorption channel can proceed and the current will flow against the applied d.c. bias. As soon as the d.c. bias exceeds $\hbar\omega/ed$ (figure 2(c)) the picture is reversed and the current will flow in the direction of the applied d.c. bias.

To model this behaviour we assume that the broadened levels have state densities $\rho(E)$ given by Gaussians, and the occupation is given by a Boltzmann distribution function, $f(E)$. The main assumption being that the carriers are distributed thermally with more at the bottom of the broadened levels than at the top. Following Tien and Gordon we weigh the current through the PAT channels with Bessel functions and model the total ground state to

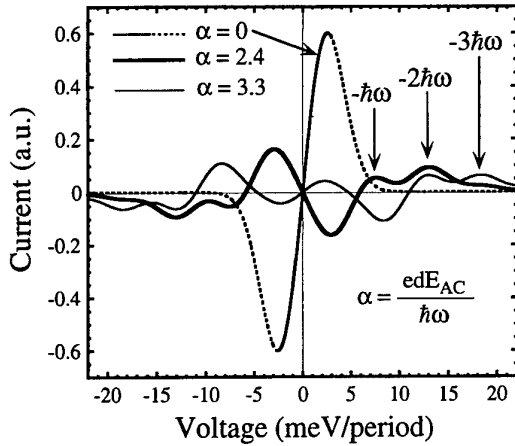


Figure 4. Current–voltage characteristics produced by the model calculation without radiation and with 1.30 THz radiation at two different laser field strengths.

ground state tunnelling current according to

$$j \propto \sum_{n=-\infty}^{\infty} J_n^2(\alpha) \int dE \rho(E) f(E) [\rho(E - eV_{DC} - n\hbar\omega) - \rho(E + eV_{DC} - n\hbar\omega)].$$

We have neglected PAT into excited states, which becomes important at high d.c. biases and large laser field strengths [27]. The Gaussian density of states is assigned a width of 2 meV, which is almost four times the calculated ground state miniband width of 0.55 meV, but is consistent with the broadened width deduced from the I – V characteristic without radiation. It is reasonable to expect that the scattering mechanisms that determine the superlattice transport properties without radiation are also important for the transport properties with THz radiation.

The result of the calculation is shown in figure 4. In the absence of the THz radiation it can be seen that the ohmic behaviour near zero bias is reproduced by the model. Beyond the ohmic region (dotted curve), however, the model begins to diverge from the experiment. This is due to the fact that the model assumes that the electric field is uniform throughout the superlattices and does not take into account the effect of electric field domains. At the first zero of $J_0^2(\alpha)$ ($\alpha = 2.4$) the zero-photon current has been shut off, the current can only flow through the photon-assisted channels and the current becomes negative near the origin. At the next higher value of laser field strength ($\alpha = 3.3$), a three-photon emission step is present. Overall, the essential features of the experimental I – V characteristic are reproduced by this simple model.

Negative current flow for small positive d.c. bias requires level broadening and tails of the energy levels to extend to within reach of the photon-assisted channel. In this view the absolute negative conductance is an unexpected but reasonable manifestation of dynamic localization. Furthermore, under d.c. bias the ladder of quantum well ground states will all be equally occupied and, at first blush, net stimulated absorption or emission of a photon should be absent [3]. By introducing scattering it

is apparent that net gain (stimulated emission) can appear on the low-frequency side of the Stark ladder splittings and net loss (stimulated absorption) can occur on the high-frequency side, relieving the apparent contradiction with Bastard *et al* [3].

In conclusion, we have observed dynamic localization in a multi-quantum well superlattice in the sequential tunnelling limit. The dynamic localization is accompanied by negative absolute conductance around zero bias followed by the appearance of photon-assisted channels that correspond to the stimulated emission of as many as three terahertz photons. The essential features of the I – V characteristics are reproduced by a simple empirical model in which virtual states play a similar role to the unperturbed quantum well states.

Acknowledgments

We would like to thank the staff at the Center for Free-Electron Laser Studies, J R Allen, D Enyeart, G Ramian and D White. Funding for the Center for Free-Electron Laser Studies is provided by the Office of Naval Research. Additional funding was provided by the Army Research Office, the NSF, and the Air Force Office of Scientific Research. BJK would also like to thank Gloria Platero for useful discussions.

References

- [1] Pavlovich V V and Epshtein E M 1976 *Sov. Phys. Semicond.* **10** 1196
- [2] Dakhnovskii Y and Metiu H 1995 *Phys. Rev. B* **51** 4193
- [3] Bastard G, Brown J A and Ferreira R 1994 *Solid State Physics: Semiconductor Heterostructures and Nanostructures* vol 44, ed H Ehrenreich and D Turnbull (San Diego, CA: Academic) p 325
- [4] Tien P K and Gordon J P 1963 *Phys. Rev.* **129** 647
- [5] Dayem A H and Martin R J 1962 *Phys. Rev. Lett.* **8** 246
- [6] Hamilton C A and Shapiro S 1970 *Phys. Rev. B* **2** 4494
- [7] Büttiker M and Landauer R 1982 *Phys. Rev. Lett.* **49** 1739; 1986 *IBM J. Res. Dev.* **30** 451
- [8] van der Heijden R W, Swartjes H M and Wyder P 1984 *Phys. Rev. B* **6** 3513
- [9] Tucker J R and Feldman M J 1985 *Rev. Mod. Phys.* **57** 1055
- [10] Takayanagi H 1990 *Physica B* **165–6** 959
- [11] Apell S P and Penn D R 1992 *Phys. Rev. B* **45** 6757
- [12] Apell S P, Penn D R and Stiles M D 1992 *Appl. Phys. Lett.* **72** 1175
- [13] Hu Q 1993 *Appl. Phys. Lett.* **62** 837
- [14] Feng S and Hu Q 1993 *Phys. Rev. B* **48** 5354
- [15] White J D and Wagner M 1993 *Phys. Rev. B* **48** 2799
- [16] Wagner M 1994 *Phys. Rev. B* **49** 16544
- [17] Iñárrrea J, Platero G and Tejedor C 1994 *Phys. Rev. B* **50** 4581
- [18] Wagner M 1995 *Phys. Rev. A* **51** 798
- [19] Wagner M 1995 *Proc. 22nd Int. Conf. on the Physics of Semiconductors* vol 2, ed D J Lockwood (Singapore: World Scientific) p 1079
- [20] Yakubo K, Feng S and Hu Q *Phys. Rev. B* unpublished
- [21] Iñárrrea J and Platero G 1996 *Europhys. Lett.* **24** 43
- [22] Aguado R, Iñárrrea J and Platero G 1996 *Phys. Rev. B* **53** 10030
- [23] Kouwenhoven L P, Jauhar S, McCormick K, Dixon D, McEuen P L, Nazaro Yu V, van der Vaart N C and Foxon C T 1994 *Phys. Rev. B* **50** 2019

- [24] Guimaraes P S S, Keay B J, Kaminski J P, Allen S J Jr, Hopkins P F, Gossard A C, Florez L T and Harbison J P 1993 *Phys. Rev. Lett.* **70** 3792
- [25] Keay B J, Guimaraes P S S, Kaminski J P, Allen S J Jr, Hopkins P F, Gossard A C, Florez L T and Harbison J P 1994 *Surf. Sci.* **305** 385
- [26] Keay B J, Allen S J Jr, Kaminski J P, Campman K L, Gossard A C, Bhattacharya U, Rodwell M J W and Galán J 1995 *Proc. 22nd Int. Conf. on the Physics of Semiconductors* vol 2 ed D J Lockwood (Singapore: World Scientific) p 1055
- [27] Keay B J, Allen S J Jr, Galán J, Kaminski J P, Campman K L, Gossard A C, Bhattacharya U and Rodwell M J W 1995 *Phys. Rev. Lett.* **75** 4098
- [28] Faist J, Capasso F, Sivco D L, Sirtori C, Hutchinson A L and Cho A Y 1994 *Science* **264** 553
- [29] Drexler H, Scott J S, Allen S J, Campman K L and Gossard A C 1995 *Appl. Phys. Lett.* **67** 2816
- [30] Kouwenhoven L P, Jauhar S, Orenstein J and McEuen P L 1994 *Phys. Rev. Lett.* **73** 3443
- [31] Blick R H, Haug R J, van der Weide D W, von Klitzing K and Eberl K 1995 *Appl. Phys. Lett.* **67** 3924
- [32] Verghese S, Wyss R A, Schäpers Th, Förster A, Rooks M J and Hu Q 1995 *Phys. Rev. B* **52** 14 834
- [33] Heitler W 1984 *The Quantum Theory of Radiation* (New York: Dover) p 140
- [34] Butcher P N and Cotter D 1990 *The Elements of Nonlinear Optics* (Cambridge: Cambridge University Press)
- [35] Klingshirn C F 1995 *Semiconductor Optics* (Berlin: Springer)
- [36] Hanna D C, Yuratich M A and Cotter D 1979 *Nonlinear Optics of Free Atoms and Molecules* (Berlin: Springer)
- [37] Liboff R L 1980 *Introductory Quantum Mechanics* (Oakland, CA: Holden-Day) p 582
- [38] Baym G 1973 *Lectures on Quantum Mechanics* (New York: Benjamin/Cummings)
- [39] Keay B J and Aversa C 1996 *Phys. Rev. B* **54** R2284
Keay B J 1995 *PhD Dissertation* University of California, Santa Barbara
- [40] Choi K K, Levine B F, Malik R J, Walker J and Bethea C G 1987 *Phys. Rev. B* **35** 4172
- [41] Kazarinov R F and Suris R A 1971 *Sov. Phys. Semicond.* **5** 707

Bandgap engineering of electro-optical modulators and switches

Paul Voisin

Laboratoire de Physique de la Matière Condensée de l'Ecole Normale Supérieure,
24 rue Lhomond, F75005 Paris, France

Abstract. Bandgap engineering of advanced optoelectronic devices is a natural and desired outcome of the basic studies of semiconductor quantum well structures. This is first illustrated with the problem of the polarization sensitivity of guided-wave devices. It is shown that tensile-strained quantum wells near the light- and heavy-hole degeneracy provide almost equal absorption spectra and nearly identical electroabsorption curves in the TE and TM modes, which are necessary features for on-line applications in optical fibre telecommunication systems. In the second part, the ultimate performances of photovoltaic bistable devices based on the Wannier–Stark effect are discussed. It is shown that the commutation energy of today's silicon MOSFETs (≈ 30 fJ) can be challenged by usable optical devices.

1. Introduction

The concept of bandgap engineering was first proposed by Capasso more than ten years ago. The idea is to capitalize on the considerable effort devoted to the fundamental electronic, optical or electro-optical properties of semiconductor heterostructures, and to make use of their extraordinary flexibility to model and actually realize innovative devices. Yet, in the past many astute ideas have missed their target because the physicists involved have neglected the constraints that the devices must obey to be useful ones. A major field of application of these concepts is optoelectronics, and more precisely the devices required for high-bit-rate optical fibre telecommunication systems. Here, we shall discuss two examples of current interest, namely guided-wave electroabsorption modulators and bistable devices for commutation matrices or other optical-logic applications. In the first case, one starts with devices which have already demonstrated their potential for high-speed modulation, but which suffer from a strong sensitivity to the polarization of the incoming light which hinders their use for in-line modulation. In this situation, the needs are well defined and the bandgap engineer's role is to find out the simplest, best performing and most reproducible solution fulfilling the numerous and precise constraints imposed by system operation. The combination of built-in strains and quantum confinement actually leads to an elegant solution of this problem. Conversely, in the commutation matrix problem, one deals with a more open field because there are many possible architectures for a commutation node, and the ideal mixing of purely optical, optoelectronic and purely electronic devices is not so well defined. Besides, there might be other applications

of optical bistable matrices, for instance in dedicated optical processors. Major parameters for a bistable switch are the power needed to maintain the device state and the commutation energy required to switch the device. The most popular approach is certainly the self-electro-optical effect device (SEED) invented by Miller in 1985. Recent advances in photovoltaic superlattice SEEDs will be discussed, with the salient result that the commutation energy of today's MOSFETs (≈ 30 fJ/gate) can actually be challenged by all-optical devices.

2. Polarization-insensitive guided-wave devices

Polarization-independent operation is a desired feature for many optoelectronic devices used in telecommunication systems [1]. Indeed, the incident polarization is randomized after propagation through optical fibres, and on-line devices have to be polarization independent. This is particularly important for electro-optical modulators and amplifiers. As device lengths of the order of ≈ 100 μm are needed, only the in-plane propagation can provide the required performances. Hence, these devices are usually waveguides, and they unfortunately present a strong birefringence, which results from the combination of a number of factors including the waveguide geometry, the optical selection rules, the transition strength and the nature of the electro-optical effect. Electroabsorption (EA) modulators are of particular interest for they are ideally suited for high-bit-rate (>10 GHz) operation, and semiconductor amplifiers are essential for long-haul transmission like intercity or submarine optical cables. Heterostructures based on III–V semiconductor multiple quantum wells (MQW) have widely proved their potential

for the above-mentioned applications, but waveguides using lattice-matched MQWs suffer from a strong sensitivity to the incident light polarization. This is inherently due to the confinement splitting of the light- and heavy-hole levels associated with the polarization selection rules acting on the heavy-hole to electron (H-E) and light-hole to electron (L-E) optical transitions, which lead to different bandgaps in the TE and TM modes. Indeed, the H-E transition dominates the absorption in the TE (or in-plane polarization) mode whereas the L-E transition is the only contribution to the absorption in the TM (or axial polarization) mode. Clearly, the first requirement to achieve polarization independence is the use of tensile strained material in the wells, in order to obtain a compensation between the strain-induced and the confinement-induced splittings and get degenerate L-E and H-E transitions. Furthermore, guided-wave absorption coefficients for the TE and TM modes have to be equal to provide identical gains in amplifiers, whereas the electric-field dependence of these absorptions must be the same for polarization-insensitive modulation. In the latter case, the nature of the electro-optical effect (for instance, the quantum-confined Stark effect [2] or the Wannier-Stark effect [3,4]) has to be considered.

We start with the guided-wave absorption coefficient for an electromagnetic wave propagating in the layer plane of the heterostructure. We consider an electromagnetic (em) guided wave characterized by an electric field $E(z) = E(z)\exp(i\omega t - k_y y)\varepsilon$, propagating along the y direction in the layer plane of the heterostructure, with the polarization ε . The calculation follows the classical procedure: the rate at which photons disappear is first calculated from Fermi's golden rule, and then related semiclassically to the absorption coefficient α , which gives [5]

$$\alpha(\omega) = \frac{E(z_0)^2}{\langle E(z)^2 \rangle} (\pi e^2 / n \varepsilon_0 c m_0^2 \omega \Omega) \times \sum_{i,f} |\langle f | \varepsilon \cdot \mathbf{p} | i \rangle|^2 \delta(E_f - E_i - \hbar\omega) \times (f(E_i) - f(E_f)) \quad (1)$$

where $\Omega = L_x L_y L_z$ is the volume in which the optical mode is normalized. If we compare equation (1) with the classical expression of the absorption coefficient for an em wave propagating parallel to the growth axis z [6], we see that the only difference is the presence of the term $E(z_0)^2/\langle E(z)^2 \rangle$.

We first assume parabolic in-plane dispersion for the valence subbands, which allows an analytical calculation of $\alpha(\omega)$. For a H(L)-E transition whose transition energy is $\varepsilon_{H(L)}$, equation (1) becomes

$$\alpha(\omega) = \frac{E(z_0)^2}{\langle E(z)^2 \rangle} \times (\beta / L_z) \mu_{H(L)} M_{H(L)} P_{H(L)} Y(\varepsilon_{H(L)} - \hbar\omega) \quad (2)$$

where $\mu_{H(L)}$ is the in-plane reduced mass of the corresponding electron-hole pair (in units of the conduction effective mass m_c), $M_{H(L)}$ the squared overlap of the electron and heavy- or light-hole envelope functions, $P_{H(L)}$ the variable part of the relevant atomic dipole matrix element, and $Y(x)$ is the step function. $\beta = e^2 E_p m_c / 4 n \varepsilon_0 c m_0 \hbar^2 \omega \approx 6 \times 10^{-3}$ is a dimensionless constant ($E_p \approx 23$ eV is Kane's matrix element). For

a multiple quantum well structure, we finally obtain the expression of the guided-wave absorption coefficient by summing over the well positions, which amounts to replacing the $E(z_0)^2/\langle E^2 \rangle$ factor in equation (2) by Γ/L_w , where the 'confinement factor' $\Gamma = \int_{\text{wells}} E(z)^2 dz / \langle E^2 \rangle$ is the overlap of the electromagnetic mode with the quantum wells. Γ is obviously proportional to the well width L_w , so the guided wave absorption does not depend on L_w . Actual modulator structures usually have about 10 wells, and calculation of the mode profile leads to Γ values of the order of 25–30%.

In the diagonal (or parabolic) model, the heavy- and light-hole in-plane and axial (i.e. along the quantization axis z) masses are given in terms of the Luttinger parameters by $m_H^I = 1/(\gamma_1 + \gamma_2)$, $m_L^I = 1/(\gamma_1 - \gamma_2)$, $m_H^z = 1/(\gamma_1 - 2\gamma_2)$, $m_L^z = 1/(\gamma_1 + 2\gamma_2)$. In InGaAs, the heavy-hole mass m_H^z is $\approx 0.4m_0$, while the Kane model gives $m_L^z = 0.057m_0$ and $m_c = 0.042m_0$. $\mu_H = 0.632m_c$ and $\mu_L = 0.791m_c$ hence differ appreciably. Conversely, the wavefunction overlaps $M_{H(L)}$ (in the absence of an external electric field) are essentially equal to unity. As for the atomic part of the optical matrix element, we have from the definition of the electron, light-hole and heavy-hole Bloch functions $P_H = 1$, $P_L = 1/3$ for the TE mode, and $P_H = 0$, $P_L = 4/3$ for the TM mode. Hence the sum of the L-E and H-E contributions in the TE mode just equals the L-E contribution in the TM mode. Using these relations, we can readily calculate the guided wave absorption spectra for unstrained or strained quantum wells. For instance, for a 13 QWs modulator with $L_w = 121$ Å and $\Gamma = 27\%$, the contribution of the H-E transition gives $\alpha = 856$ cm⁻¹. α is conveniently expressed in logarithmic units, using a reference waveguide length. 856 cm⁻¹ corresponds to a modal attenuation of 37 dB/100 μm.

The diagonal model discussed up to now is particularly convenient for it leads to analytical expressions and is a very good guide-line, but it is not quite satisfactory: it is now well known that when the in-plane vector is non-zero, the heavy- and light-hole states become coupled and anticrossings between heavy- and light-hole branches occur which eventually lead to camel-back-shaped dispersion relations. Figures 1(a) and 1(b) show the comparison between essentially exact in-plane dispersion relations (broken curves) and the predictions of the diagonal model (full curves), in the representative case of 121 Å wide InGaAs-AlInAs QWs. Figure 1(a) is for unstrained quantum wells and figure 1(b) for the case of tensile strained quantum wells at the exact heavy- and light-hole degeneracy. The anticrossings correspond to a strong admixture of light- and heavy-hole characters in the eigenstates. It is clear that the exact in-plane dispersion of the valence subbands has to be considered for a proper evaluation of the TE and TM absorption coefficients. The calculated absorption spectra are displayed in figures 1(c) and 1(d), the upper part showing the prediction of the diagonal model and the lower part the result of the exact calculation. The TE/TM sensitivity of the confinement factor Γ is not included at this point and will be discussed later on. In the unstrained case, the spectra are in both cases dominated by the expected bandgap difference between

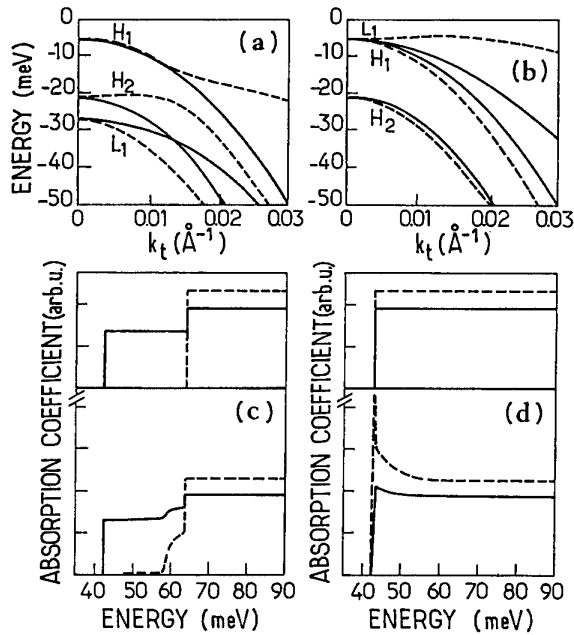


Figure 1. In-plane dispersion relations of the valence subbands (a) without strain and (b) with the biaxial tensile strain compensating the confinement splitting of L1 and H1. The full curves are the subband dispersions in the parabolic model and the broken curves show the exact dispersions. Absorption coefficients of the unstrained (c) and strained (d) quantum well in the parabolic model (up) and with the exact in-plane dispersion (down). The full (broken) curves correspond to the TE (TM) mode. Calculations are for a 121 Å thick GaInAs–AlInAs QW.

the TE (full curves) and the TM modes (broken curves). Band mixing effects simply produce a small splitting of the L–E transition, due to the strong interaction between the neighbouring level L1 and H2. For the tensile-strained case, the absorption spectrum is a single-step structure showing equal TE and TM absorption gaps. The camel-back shape of the light-hole dispersion results in a spike close to the absorption edge, which is of course more pronounced in the TM mode. However, this non-parabolicity effect appears only in a narrow spectral range, and broadening effects (≈ 7 meV for best materials at 300 K) just smear the spike out. Independent on the model, it can be observed that a finite difference remains between the TE and TM absorptions, the later being slightly ($\approx 20\%$) more effective. This is an unavoidable consequence of the different joint density of states for the H–E and L–E transitions.

We now discuss the TE–TM sensitivity of the confinement factor Γ , which is due to the different boundary conditions for the electromagnetic field. For this we consider the model of a three-layer slab waveguide formed by a uniform isotropic layer of refractive index n [7], sandwiched between two semi-infinite layers of lower refractive index. Again, the z -axis is taken perpendicular to the layers and the light propagation direction along the y -axis. For guided modes, the requirement is to have an oscillatory solution in the core with evanescent tails in the cladding regions. The TE and TM configurations

correspond respectively to $E(z) = E_x(z)x$ (TE mode) and $H(z) = H_x(z)x$ (TM mode), where x is the unit vector in the x direction. From the Helmholtz wave equation associated with the appropriate boundary conditions at each interface, we find the expressions of the components of the E and H vectors in the three layers for both modes as functions of the longitudinal propagation constant and the guide geometry. The propagation constant is calculated by using the effective index method. The confinement factor Γ is obviously the ratio of the guided light power in the wells by the total power. To calculate it, we calculate the Poynting vector flow $\Pi = L_x \int \pi_y dz$, using $\pi_y = \frac{1}{2} E_x H_z$ for the TE mode and $\pi_y = \frac{1}{2} H_x E_z$ for the TM mode respectively. Calculations for a structure with 13 InGaAs 121 Å wide QWs separated by AlInAs 80 Å barriers and cladded by AlInAs layers, using standard values of the refractive indices yield $\Gamma^{\text{TE}} = 27\%$ and $\Gamma^{\text{TM}} = 23\%$. It is noteworthy that this $\approx 20\%$ difference is not very sensitive to the exact waveguide structure. Almost the same figure is obtained with the parameters of a GaInAs–GaInAsP MQW cladded between InP layers.

Quite remarkably, the polarization dependence of Γ thus tends to compensate the effect of the joint density of states. We thus expect nearly identical guided-wave absorption coefficients in the TE and TM modes (in the flat-band regime) in heterostructures using tensile-strained MQW structures at the heavy- and light-hole degeneracy. This condition is sufficient to ensure identical TE and TM gains for optical amplifiers at gain saturation. However, to obtain a polarization-independent modulation, the field dependence of the absorption must also be the same in both modes. The dominant electro-optical effect in the structures under investigation is the quantum confined Stark effect (QCSE). The field-induced variations of the H–E and L–E transitions have to be considered both in terms of the Stark-shift and wavefunctions overlaps $M_{H(L)}$.

First, as illustrated in figure 2(a), the Stark shift in a square quantum well depends on the carrier effective mass, and differs appreciably for the H–E and L–E transitions. But at the same time, M_H decreases with the applied field faster than M_L , as shown in figure 2(b). Again, we expect kind of a compensation between these two effects: in the TE mode the absorption edge shifts faster with the applied field than in the TM mode but its strength also decreases faster. A realistic calculation of the electroabsorption curves requires the introduction of a finite broadening of the transitions. In figure 3 we show calculated electroabsorption curves obtained for realistic device parameters, i.e. a working wavelength 15 meV below the MQW bandgap, a spectral line broadening $\gamma = 7$ meV (we use a Gaussian broadening of the optical transitions), and a 13 QW GaInAs–AlInAs MQW, with $L_w = 121$ Å. It can be seen that in spite of the large difference between the heavy- and light-hole Stark shifts, the TE and TM electroabsorptions are nearly identical, from transparency down to a very large attenuation of 20 dB/100 μm . The main difference is that the TM absorption maximum is larger and obtained for a larger electric field than the TE one. Let us point that this difference is not significant for a device length above

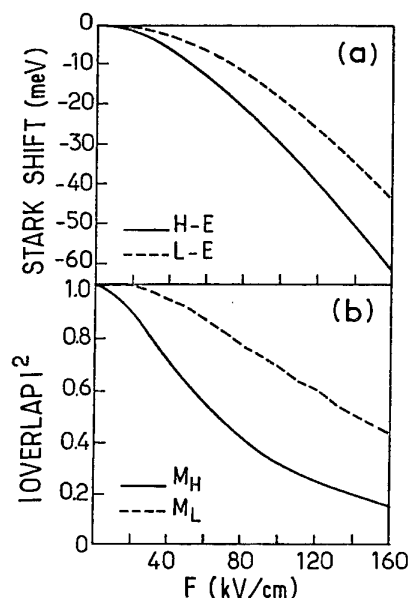


Figure 2. Stark shifts (a) and envelope functions overlaps $M_{H(L)}$ (b) for the H1-E1 (full curves) and L1-E1 (broken curves) transitions in a 121 Å GaInAs-AlInAs QW.

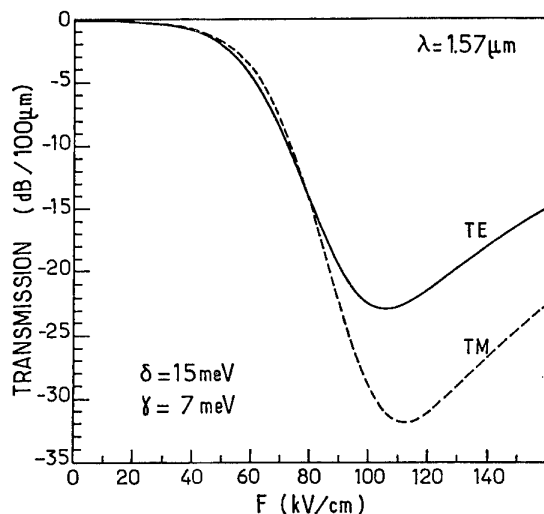


Figure 3. Calculated electroabsorption curves for a 13 QWs (121 Å thick) GaInAs-AlInAs MQW, including a spectral broadening $\gamma = 7 \text{ meV}$ of the optical transitions.

100 μm , for it appears in a regime where the device transmission is below 1%.

Finally, it should be stressed that this analysis does not take into account excitonic corrections. It is clear that exciton bound states do not play a major role in QWs based on GaInAs, due to their modest binding energy. However, Coulomb enhancement does exist and is expected to increase the band-to-band absorption coefficient by about 50% [8]. This enhancement should be nearly polarization independent.

To conclude this section, it appears that tensile-strained MQW structures at the heavy- and light-hole degeneracy

are appropriate to the realization of essential optoelectronic devices, namely polarization-insensitive amplifiers and EA modulators. Polarization insensitivity is not an exact result, but stems from a quasi perfect compensation of different contributions. Note that these theoretical results are in close agreement with recent experimental findings in the GaInAs-InP [9] and GaInAs-AlInAs [10] families of materials. Clearly, this detailed modelling and its comparison with experiments opens the way for the final optimization of these devices.

3. Photovoltaic optical bistable devices

There is also a strong interest in optical bistable devices because of their potential applications in commutation matrices for telecommunications, parallel information processing, etc. Electro-optical bistable device like the self-electro-optical effect device (SEED) [11] have received considerable attention because they function at low power (a few microwatts per device), but they need an individual electrical circuit, which implies a complex architecture of interconnections and limits the possibility of large-scale integration. Recently, the invention of the S-SEED [12] has opened new perspectives since this combination of two SEEDs functions with no other external electrical connection than a simple bias voltage. However, these devices still require a differential illumination which hamper their use in large-scale matrices. All-optical bistability using optical non-linearities in Fabry-Pérot cavities [13] remains an attractive alternative and is also receiving much attention [14]. Here, we discuss an other direction which is biasless (or even wireless) photovoltaic SEEDs using the combination of the Wannier-Stark effect, built-in electric field and screening [15].

The experiments are carried out on a p-i-n diode structure which was originally designed for guided-wave modulation at 1.55 μm [4]. The core consists of a 20 period InGaAs (60 Å)-InAlAs (15 Å) superlattice inserted between two layers of 10-period InGaAs (25 Å)-InAlAs (20 Å). This undoped structure is bounded by thick InAlAs n+ and p+ electrodes and finally topped by a 1000 Å p+ InGaAs contact layer. This sample was grown by MBE on an n+ InP substrate. We used $\approx 1 \times 2 \text{ mm}^2$ pieces of the unprocessed sample. The experimental set-up uses a simple halogen lamp followed by a spectrometer as the tunable source, and InGaAs photodiodes for detection. Most experiments were done under cw conditions. A simple silver-glue electrical contact is sometimes added and allows simultaneous measurements of the electrical and optical properties. Deep-etched 100 μm mesa photodiodes have also been processed and examined.

Figure 4 shows photocurrent spectra obtained at 77 K with a forward bias V_f from 0 V to 0.7 V. These spectra illustrate the now well known Wannier-Stark effect [3, 4, 16]: in the presence of an electric field, the SL optical absorption spectrum $\alpha(h\nu)$ consists in a series of step-like transitions at the energies $E_0 + peFd$, where E_0 is the bandgap of an isolated quantum well, F is the electric field, d the superperiod and p a relative integer. The sample used in the experiments has a built-in electric

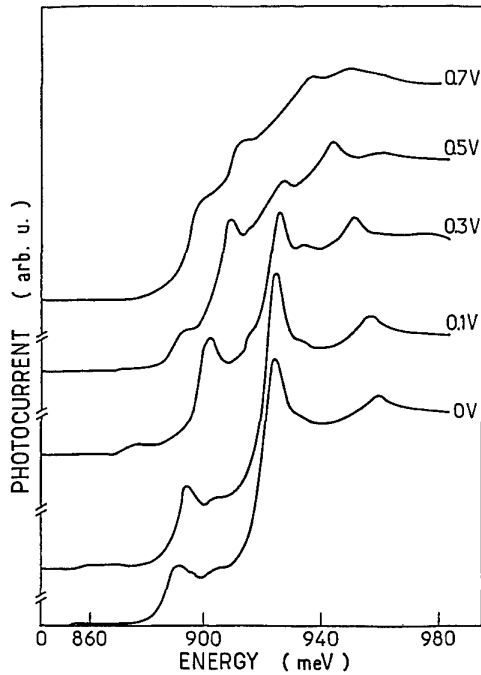


Figure 4. 77 K photocurrent spectra at various forward bias voltages in a 60 Å–20 Å GaInAs–AlInAs superlattice.

field of 50 kV cm^{-1} which already corresponds to the 'high field' regime of the Wannier–Stark effect. At $V_f = 0 \text{ V}$, only the dominant 'vertical' transition $p = 0$ at 920 meV (corresponding to electron and hole wavefunctions centred in the same quantum well) and the first $p = \pm 1$ 'oblique' transitions at 890 and 960 meV (corresponding to electron and hole wavefunctions localized around adjacent quantum wells) are observed. The low-energy oblique transition is the effect used for guided-wave optical modulation [4]. Applying a forward bias decreases the electric field across the SL: the $p = -1$ transition is blue-shifted and gains oscillator strength, and additional $p = -2$ and $p = -3$ transitions appear and develop. At $V_f = 0.7 \text{ V}$, the built-in field is almost perfectly screened and the sample is in near flatband condition: the wavefunctions are delocalized over the superlattice and a miniband absorption spectrum appears. These spectra show around 908 meV a region of contrasted 'blueshift' (or negative electroabsorption) modulation, where the absorption increases when the field decreases.

The general principle of optical bistability due to electro-optical feedback [11] is the screening of the built-in (or applied) electric field by photocarriers, in a spectral region of negative electroabsorption. Raising the incident power increases the number of photocarriers. They tend to screen the electric field, which leads to an increase of the absorption. Optical absorption hence causes its own increase. If the absorption change is sharp enough, the sample can switch hysteretically between its weakly absorbing and highly absorbing regimes. This behaviour is illustrated in figure 5, which shows the transmitted power P_{out} as a function of the incident power P_{in} observed in our sample for various working wavelengths, at a temperature

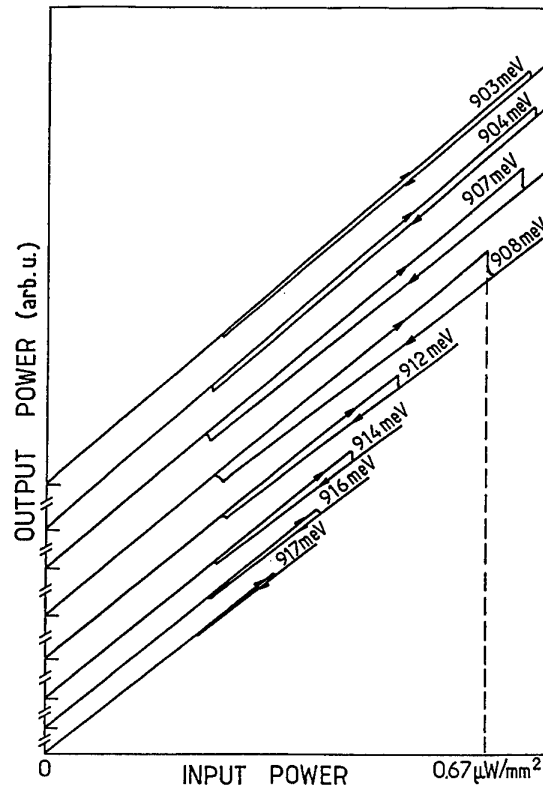


Figure 5. Transmitted versus incident intensities at various photon energies at 77 K.

of 77 K. The hysteresis loop is observed over a large (20 meV) spectral range. The two slopes associated with the corresponding absorption regimes are clearly observed in the hysteresis loop at 908 meV. The incident power at the centre of the loop is less than $1 \mu\text{W mm}^{-2}$. Moreover, the 2:1 ratio between the switching-up and switching-down powers $P_{\text{in}\uparrow}$ and $P_{\text{in}\downarrow}$ offers a nice insensitivity to input power noise.

Although the way this device works presents strong analogies with the SEED, the major difference is that no load circuit is needed. In analogy with SEEDs, the system status is determined by the intersections of the electroabsorption curve $\alpha(V)$ and a 'load line', which is obtained by writing that the sum of the photocurrent and the diffusion-drift current is zero. Experimentally, the latter is given by the sample $I(V)$ characteristic, which is shown in figure 6(a). The $I(V)$ curve shows a spectacular negative dynamical resistance (NDR) effect. The origin of this NDR is not definitely established. Tunnelling between the SL and the optically passive region surrounding it is likely to play the leading role. The photocurrent I_{ph} is proportional to the absorbed power $(1 - e^{-\alpha})P_{\text{in}}$. Here, the maximum absorption is ≤ 0.1 , so we can drop the exponential and get the load line

$$\alpha = I(V)/\beta P_{\text{in}}. \quad (3)$$

The $\alpha(V)$ curve at 908 meV, measured from the combination of transmission and photocurrent data, is plotted in figure 6(a) together with the load lines (equation (3)), obtained for increasing values of βP_{in} .

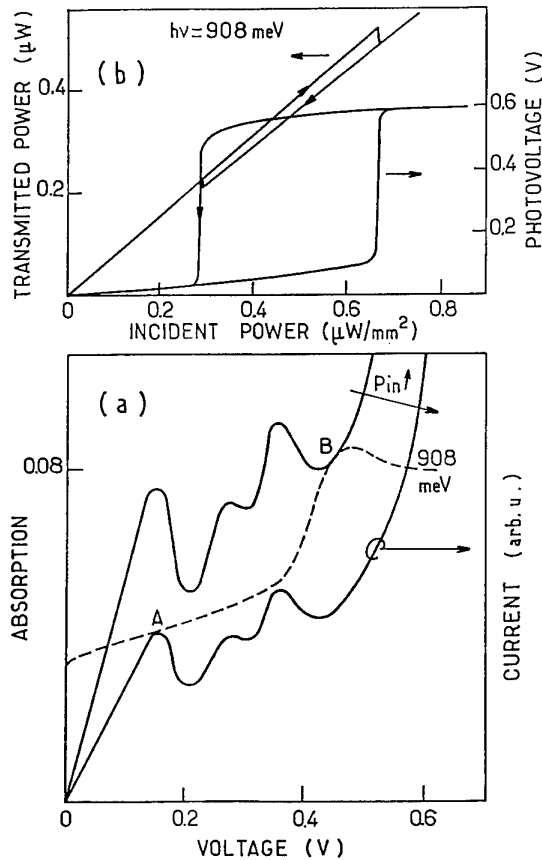


Figure 6. (a) Absorption versus voltage curve for $h\nu = 908$ meV (broken curve), and 'load lines' (equation (3)) for increasing values of the incident power P_{in} (full curves). The load lines are proportional to the measured $I(V)$ characteristic (right vertical scale), and the low-voltage sample resistance is $2\text{ M}\Omega$. (b) Photovoltage and transmitted power versus incident power at 908 meV.

It is clear that there is a domain of βP_{in} where three intersections exist. An argument quite similar to the original discussion of Miller *et al* proves that the central point is unstable, while the two others correspond to the stable states of the hysteresis loop. It is seen at first glance that the contrast in transmitted power P_{out} is mainly due to the electroabsorption curve, while the contrast in P_{in} thresholds is mainly a consequence of the strong NDR: drawing a straight line instead of the actual load line would obviously lead to a much narrower hysteresis loop. The quantitative agreement between this simple model and the observed optical hysteresis can be checked in more detail: the switching-up (-down) threshold $P_{in\uparrow}$ ($P_{in\downarrow}$) corresponds to the load line reaching the point 'A' ('B') in figure 6(a), which yields a 2:1 P_{in} ratio. Conversely, the 2:1 absorption difference observed in figures 4 and 6(a) does correspond to the 6.5% optical transmission contrast observed in figures 2 and 6(b), if we take into account that the measured absorption on the two-dimensional plateau (above 980 meV in figure 4) is 20%. In addition, we can measure simultaneously the optical transmission and photovoltage bistabilities, as shown in figure 6(b). The threshold photovoltages observed in figure 6(b) (0.17 V and

0.4 V) are again in excellent agreement with the positions of the 'A' and 'B' points in figure 6(a).

The transmission contrast decreases with increasing temperature, and ultimately disappears near 215 K. This results mainly from the smoothing of the electro-optical response, which diminishes the contrast between the 'weakly' and 'highly' absorbing regimes. It is noteworthy that the switching-up power, into a higher absorption regime, remains constant with temperature: activation of carrier diffusion does not seem to affect this device near room temperature. In other words, the sample peak-current associated with point 'A' is essentially temperature independent. On the contrary, the switching-down power tends to increase with temperature, which indicates that the valley current corresponding to point 'B' shows some thermoactivation. These observations are consistent with a tunnel-diode origin for the observed NDR, but do not help the elucidation of its origin. In fact, while remarkable and stimulating, the strong undesigned NDR observed in this sample is not necessary to the bistability. Actually, it disappears in other parts of the wafer where the transition energies are shifted by a few meV, due to small lateral inhomogeneities of the epitaxial growth. In small photodiodes processed in this part of the wafer, well contrasted but narrower hysteresis loops are observed and fitted with a regular diode $I(V)$ curve for the load line (equation (3)). Finally, let us note that designing a sample having a larger built-in electric field would decrease the initial strength of the oblique transition and improve the absorption contrast. This is likely enough to ensure room-temperature operation. Indeed, we have been able to restore the bistability loop at room temperature by adding a reverse bias of -1.5 V, which simply simulates a larger built-in electric field.

The working power $P < 1\text{ }\mu\text{W mm}^{-2}$ is clearly outstanding, but the parameter of ultimate interest for a bistable device is the switching energy E_s . To measure E_s , the time dependence of the effect must be investigated. For this, we merely chop the incident beam and measure directly the photovoltaic signal on an oscilloscope having a $1\text{ M}\Omega$ input resistance somewhat smaller than, but comparable to, the 1 mm^2 sample low-voltage resistance. The excitation power is of the order of $P_{in\uparrow}$. We find exponential-like responses with a characteristic time $\tau \approx 3$ ms. Using $\Delta P_{in} = 0.5 \times 10^{-6}\text{ }\mu\text{W }\mu\text{m}^{-2}$, we get a switching energy E_s as small as $2\text{ fJ }\mu\text{m}^{-2}$! This value compares favourably with the 30 fJ/gate switching energy of today's silicon MOSFETs. Besides, the conservation of energy gives us another way to estimate the switching energy: from the electrostatic point of view, the device is essentially a capacitor, and E_s is nothing other than the corresponding change in the electrostatic energy stored in the capacitor. Counting only the absorbed power αP_{in} , gives $E_s = \frac{1}{2}CV^2/\alpha$. Using $V = 0.6$ V, we get an ultimate value $E_s = 1.8\text{ fJ }\mu\text{m}^{-2}$ in close agreement with our experimental result. In addition, it is noteworthy that the relaxation time τ is essentially an RC time constant where $R = 1\text{ M}\Omega$ is the impedance of the oscilloscope, and C is the device capacitance. From the sample parameters, we have $C = 1\text{ nF mm}^{-2}$, which gives $\tau \approx 2$ ms, in

fair agreement with the measured 3 ms. This RC time can be reduced by shunting the device with a smaller resistance. However, the optical power necessary to screen the built-in electric field increases proportionally. Note also that the switching-up time is ultimately associated with the photocarrier drift through the superlattice, while the switching-down time is rather governed by the external resistance. A clear difference is actually observed between the fast rise time and slow decay time of the photovoltage signal.

Finally, the consideration of the load line shape in figure 6(a) is an invitation to imagine a completely new type of electro-optical bistability relying exclusively on negative dynamical resistance and not on the positive feedback associated with negative electroabsorption. This effect is actually observed at the photon energy $h\nu = 892$ meV, where the sample switches from an absorbing state to a more transparent one when increasing P_{in} , which is at odds with the SEED bistability. The interpretation is the following: as can be observed in figure 4, $h\nu = 892$ meV corresponds to the spectral region below the superlattice bandgap, where large-field electroabsorption is due to the 'oblique transition' $p = -1$. When the field decreases, this transition is blue shifted, and the sample becomes nearly transparent: we have a positive electroabsorption increasing with increasing electric field, i.e. decreasing with increasing forward bias. This usually leads to monostable operation. However, in the presence of a strong NDR, multistability is still possible, because the oscillation in the load line overcompensates the absorption decrease. Although the bistable loop is smaller than at 908 meV, this new 'antiSEED' effect is of particular interest for it is 'complementary' to the classical one in more or less the same sense as electron-channel and hole-channel field effect transistors are 'complementary'. In addition, it should be stressed that the absorption contrast is as good as or even slightly better than at 908 meV.

To conclude this section, the high sensitivity of the Wannier-Stark effect allows the observation of electro-optical bistability with no applied bias, which leads to a record switching energy as small as $2 \text{ fJ } \mu\text{m}^{-2}$. In addition, we have observed novel all-optical bistability effects based on a strong built-in negative dynamical resistance. Depending on the sign of the electroabsorption effect, this NDR may cause two qualitatively different, in some respect complementary, bistabilities. The temperature dependence of the effects leaves us hope that modest improvement of the structure will allow room-temperature operation. Let us stress that the transmission contrast of such a structure is necessarily small (a few per cent), because the number of periods has to remain small in order to ensure the 'high-field' regime with no bias. However, the absorption contrast is large, and inserting the structure in a Fabry-Pérot cavity is a straightforward solution to the modulation contrast problem. Finally, it is noteworthy that the working wavelength here is in the $1.5 \mu\text{m}$ wavelength range, directly suitable for telecommunication applications, but the Wannier-Stark effects may be implemented in other materials as well.

4. Conclusions

We have tried to illustrate a 'basic' approach to device physics. This approach, based on detailed characterization and modelling of the properties, contrasts with the more empirical approach generally followed in industrial laboratories. Whether this time- and energy-consuming comprehensive approach is more or less efficient than the simple parametrical optimization certainly depends on the problem under consideration and on the degree of maturity of the technology. It seems obvious that complex problems with interfering parametric dependences should benefit from a detailed analysis. Conversely, one should always keep in mind that detailed modelling neglecting a single relevant issue will always miss its target. In this regard, complete comparison of modelling and experimental results is essential.

Acknowledgments

LPMC-ENS is Unité de Recherche Associée at CNRS. The contributions discussed in this paper formed the core of S Chelles' and of J Couturier's Doctorat Theses. Close collaboration with epitaxists and device physicists at CNET-Bagneux is gratefully acknowledged.

References

- [1] Suzuki M, Tanaka H, Edagawa N and Matsushima Y 1992 *J. Lightwave Technol.* **10** 1912
- [2] Miller D A B, Chemla D S, Damen T C, Gossard A C, Wiegmann W, Wood T H and Burrus C A 1985 *Phys. Rev. B* **32** 1043
- [3] Bleuse J, Bastard G and Voisin P 1988 *Phys. Rev. Lett.* **60** 220
- [4] Bigan E, Allovon M, Carré M and Voisin P 1990 *Appl. Phys. Lett.* **57** 327
- [5] Chelles S, Ferreira R and Voisin P 1995 *Semicond. Sci. Technol.* **10** 105
- [6] Bastard G 1988 *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Ulis: Editions de Physique)
- [7] Adams M J (ed) 1981 *An Introduction to Optical Waveguides* (New York: Wiley)
- [8] Lefebvre P, Christol P and Mathieu H 1993 *Phys. Rev. B* **48** 17308
- [9] Chelles S, Ferreira R, Voisin P, Ougazzaden A, Allovon M and Carencu A 1994 *Appl. Phys. Lett.* **64** 3530
- [10] Chelles S, Ferreira R, Harmand J C and Voisin P 1995 *Appl. Phys. Lett.* **67** 247
- [11] Miller D A B, Chemla D S, Damen T C, Wood T H, Burrus C A, Gossard A C and Wiegmann W 1985 *IEEE J. Quantum Electron.* **21** 1462
- [12] Lentine A, Hinton H S, Miller D A B, Henry J E, Cunningham J E and Chirovski L M F 1989 *IEEE J. Quantum Electron.* **25** 1928
- [13] Gibbs H M 1985 *Optical Bistability: Controlling Light by Light* (New York: Academic)
- [14] Oudar J L, Kuszelewicz R, Sfez B, Michel J C and Paniel R 1992 *Opt. Quantum Electron.* **24** S193
- [15] Couturier J, Voisin P and Harmand J C 1995 *Semicond. Sci. Technol.* **10** 881
- [16] Mendez E E, Agullo-Rueda F and Hong J M 1988 *Phys. Rev. Lett.* **60** 2426

Effect of donor impurities on far-infrared magnetospectroscopy of electrons in quasi-two-dimensional systems

Z X Jiang, S R Ryu and B D McCombe

Department of Physics, SUNY at Buffalo, Buffalo, NY 14260, USA

Abstract. Results of recent far-infrared magnetospectroscopic studies of a series of GaAs/AlGaAs multiple-quantum-well samples with Si donors δ -doped both in the 20 nm well centres ($2 \times 10^{10} \text{ cm}^{-2}$) and in the 60 nm barrier centres $((3.5\text{--}28) \times 10^{10} \text{ cm}^{-2})$ are described. In the presence of excess electrons a strong singlet transition of the D^- ions is observed, and one of the two predicted triplet transitions has been identified. At higher electron densities these transitions evolve continuously into blue-shifted singlet-like and triplet-like bound magnetoplasmon transitions. For the two highest-density samples the singlet-like transition exhibits clear discontinuous changes in slope versus field at integral Landau level filling factors (FFs). Plots of the difference in energy between the many-electron singlet-like transition and the D^- singlet transition normalized by the high-field Coulomb energy versus FF, exhibit cusp-like behaviour at FFs 1 and 2 at 4.2 K for the highest density sample, and at 1 and 4/3 at 1.7 K for the lower density sample. The results at integral FFs are compared with recent theoretical calculations. These experiments indicate that controlled doping of a random array of donor ions can be used to probe the nature of many-electron states in confined systems.

Many-electron effects in low-dimensional semiconductor structures, which give rise to a number of fascinating phenomena such as the fractional quantum Hall effect (FQHE), Wigner crystallization and recently the possibility of skyrmion states (Brey *et al* 1995), have been of interest for many years. Effect of these and related phenomena have been probed by interband optical techniques, but direct studies via intraband magneto-optical transitions have been largely frustrated by the fact that for a translationally invariant system the free carrier transition in a magnetic field, cyclotron resonance (CR), is sensitive only to the centre of mass motion of the entire electron gas, and not to the relative degrees of freedom (Kohn 1961). On the other hand, the simplest example of a many-electron system, a hydrogenic negative donor ion, D^- (the semiconductor analogue of the negative hydrogen ion of atomic physics), has been studied directly by FIR magnetospectroscopy for a number of years in bulk semiconductors. Recently, stable D^- ions in the absence of photoexcitation have been identified (Huant *et al* 1991) in GaAs/AlGaAs quantum wells, and additional work has verified and extended these studies (Holmes *et al* 1992, Mueller *et al* 1992, Mandray *et al* 1992). Later studies (Cheng *et al* 1993), in which excess electrons were introduced into the wells by stepwise modulation doping, exhibited interesting effects, the most striking of which were a blue-shift of the D^-

singlet transition with increasing excess electron density and discontinuities in the slope of the transition energy versus magnetic field at integral Landau level filling factors (FFs). Many-body calculations in the high magnetic field, strictly two-dimensional (2D) limit, have provided a qualitative explanation for the blue shift at FFs 1 and 2 in terms of collective transitions of the many-electron system in the presence of a positively charged impurity ion (bound magnetoplasmons) (Dzyubenko and Lozovik 1993, Hawrylak 1994). Thus GaAs/AlGaAs quantum wells with a controlled density of randomly distributed impurities in the wells provide a possible tool for probing many-body interactions of a quasi-2D electron gas, which are rendered visible in the FIR magneto-optical spectrum due to removal of translational invariance in the plane.

In order to provide a basis for understanding possible many-electron transitions, it is useful to examine the nature of the quantized single-particle states in high magnetic fields (Johnson and Lippmann 1949) in a strictly 2D situation. The energy states of an electron of effective mass m^* confined in a 2D plane perpendicular to a constant external magnetic field ($B \parallel z$ for specificity) are the so-called Landau levels, $E_N = \hbar\omega_c(N + 1/2)$, where $N = 0, 1, 2, \dots$ is the Landau level quantum number and $\omega_c = eB/m^*c$ is the cyclotron frequency. The energy states are highly degenerate, and a number of operators that commute

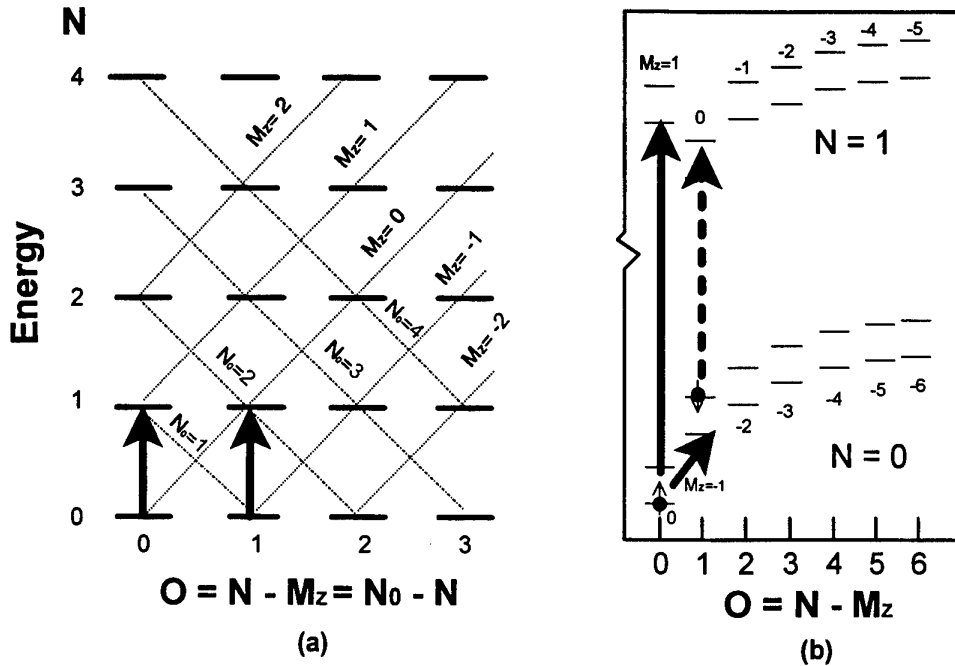


Figure 1. (a) Schematic energy level diagram showing the relationship among the various quantum numbers described in the text. Typical cyclotron resonance (CR) transitions are indicated by the bold vertical arrows. (b) Schematic single-particle energy level diagram for the two lowest Landau levels in the presence of a positive impurity ion at the coordinate origin. Two allowed transitions from the ground state are depicted as long and short full arrows ($1s-2p_{\pm}$ in the low-field, hydrogenic notation) respectively. An allowed excited state transition is shown by the broken arrow ($2p-2s$ in the low-field hydrogenic notation).

with the Hamiltonian can be defined and pairs of quantum numbers specified to determine the quantum mechanical problem completely. Two that are particularly useful are the operator that corresponds to the square of the radius of the centre of the cyclotron orbit, $r_0^2 = x_0^2 + y_0^2$, where x_0 and y_0 are operators for the coordinates of the orbit centre; and the operator corresponding to the square of the radius of the cyclotron orbit about the orbit centre, $r_c^2 = (4/m^*\omega_c^2)\hbar$, where H is the Hamiltonian. The eigenvalues of these operators are $\ell^2(2O+1)$ ($O = 0, 1, 2, 3, \dots$ is the orbit-centre quantum number) and $\ell^2(2N+1)$ respectively, where $\ell = (\hbar e/cB)^{1/2}$. The quantum numbers O and N form a convenient representation specifying the orbit-centre positions and the energy states. For each Landau level one allowed state corresponds to the area between adjacent circles of radius r_0 , $A = 2\pi\ell^2$, and the Landau level degeneracy is $\eta = 1/2\pi\ell^2 = eB/hc$; the filling factor (FF) is then $n_s/\eta = 2\pi\ell^2 n_s = hc n_s/eB$, where n_s is the 2D electron density. The operator for the z -component of angular momentum, L_z , also commutes with the Hamiltonian and has eigenvalues $\hbar M_z = 0, \pm\hbar, \pm 2\hbar, \pm 3\hbar \dots$ ($M_z \leq N$). M_z is not an independent quantum number, $M_z = N - O$. The quantity $|M_z|$ is proportional to the average distance between the electron and the origin, thus it is useful in understanding physically how an impurity located at the origin perturbs the Landau levels.

The Hamiltonian can also be decomposed into two terms, a 2D simple harmonic oscillator and $(\hbar\omega_c/2)L_z$. The oscillator quantum number, N_0 , and M_z are good quantum

numbers for this problem; $N_0 = N + O$. The parity of the wavefunctions is determined by $N + O$, $(-1)^{N+O}$. (Note that $M_z = N + O - 2O$, and therefore $(-1)^{N+O} = (-1)^{M_z}$, so M_z also independently determines the parity.) From conservation of parity and the z -component of angular momentum the electric-dipole selection rules are $\Delta M_z = \pm 1$ and $\Delta O = 0$ ($\Delta N = \pm 1$ follows); free carrier transitions between states with different orbit centres are forbidden. The energy states, quantum numbers and selection rules are summarized in figure 1(a).

Figure 1(b) shows schematically the single-particle energy levels and allowed transitions associated with the two lowest Landau levels in the presence of a Coulomb potential due to a positive charge at the origin. Note that the selection rule on O is relaxed at finite fields due to the impurity Coulomb potential, and $\Delta O \neq 0$ transitions are weakly allowed. An allowed excited state transition, $|01\rangle \rightarrow |11\rangle$, is shown by the broken arrow; this is the $2p_- - 2s$ transition in the low-field hydrogenic notation. The 'peculiar' behaviour of the $N = 1, O = 1$ energy level results from the fact that although the orbit centre is not at the origin, the probability density is peaked at the origin as indicated by $|M_z| = 0$.

The negative donor ion two-electron states and allowed transitions are straightforward to envisage. The two-electron selection rules are: $\Delta M_z^{\text{tot}} = \pm 1$; $\Delta O^{\text{tot}} \neq 0$ weak, where $M_z^{\text{tot}} = M_{z1} + M_{z2}$ and $O^{\text{tot}} = O_1 + O_2$. The basis ground state ($N^{\text{tot}} = 0, O^{\text{tot}} = 0, M_z^{\text{tot}} = 0$) is the singlet combination of spin functions with the symmetric

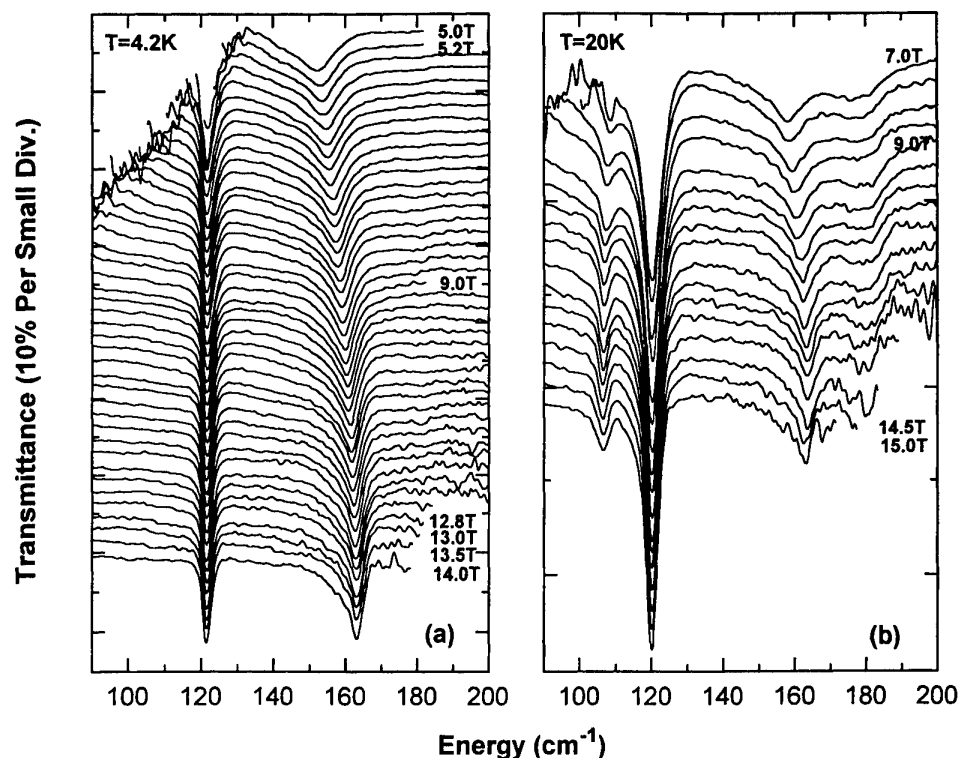


Figure 2. Magnetotransmission spectra for sample 1. (a) Data taken at 4.2 K in magnetic fields between 5 T and 13 T in steps of 0.2 T, and at 13.5 T and 14 T. (b) Data taken at 20 K in magnetic fields between 7 T and 15 T in steps of 0.5 T.

linear combination of the product $|00\rangle_{1(2)}|00\rangle_{2(1)}$ (in the notation $|NO\rangle_i$, where $i = 1$ or 2). The two strongest singlet transitions (Dzyubenko 1994) at intermediate fields terminate on excited states associated with the $N = 1$ Landau level (symmetric linear combination of the product $|00\rangle_{1(2)}|10\rangle_{2(1)}$; $N^{\text{tot}} = 1$; $O^{\text{tot}} = 0$; $M_z^{\text{tot}} = 1$), and the $N = 0$ Landau level (symmetric linear combination of the product $|00\rangle_{1(2)}|01\rangle_{2(1)}$; $N^{\text{tot}} = 0$; $O^{\text{tot}} = 1$; $M_z^{\text{tot}} = -1$). These transitions are red-shifted from their neutral donor single-particle counterparts, $|00\rangle - |10\rangle$ and $|00\rangle - |01\rangle$ respectively, due to the electron-electron repulsion, which is stronger in the ground state than the excited state. Similar considerations apply to the triplet states with appropriate antisymmetric linear combinations of the products of orbital states with the three triplet spin combinations. The antisymmetric orbital functions maximize the average separation between the two electrons, and thus minimize the repulsive Coulomb energies, leading in general to smaller shifts from their single-particle counterparts than the singlet transitions. Two strong transitions are allowed to excited spin-triplet states, both of which are associated with the $N = 1$ Landau level and track along with cyclotron resonance in the high-field limit. One transition (T_-) lies below CR (close to the neutral-donor, single-electron $|01\rangle - |11\rangle$ transition, while the other (T_+) occurs at energies higher than CR, close to the corresponding neutral-donor, single-electron $|00\rangle - |10\rangle$ transition (see figure 1(b)).

As additional excess electrons are added, exchange and correlation effects for all the electrons must be taken

into account, as well as final state interactions for optical transitions, but the dominant many-electron transitions, singlet- and triplet-like transitions, still retain the basic characters of the D^- singlet and triplet transitions outlined above. Hawrylak (1994) and Dzyubenko and Lozovik (1993) have independently calculated the many-electron transitions in the presence of a positively charged donor ion in the strictly two-dimensional, high-magnetic-field limit at integral filling factors. Both obtain a blue-shift of the many-electron transitions compared with their two-electron counterparts. The physical origin of the shift is largely exchange lowering of the many-electron ground state relative to the excited state; the shift is thus larger at $FF = 2$ than at $FF = 1$.

The far-infrared magnetospectroscopy studies were primarily carried out with a BOMEM DA-3 Fourier transform infrared spectrometer in conjunction with 9 T and 17 T superconducting magnetic systems, light-pipe, condensing cone optics, and Ge:Ga photoconductive detectors. Measurements were made at temperatures between 1.7 K and 40 K. The three MBE samples investigated were part of a larger set (see, for example Cheng *et al* 1993) all having the same basic structure, 20 nm GaAs wells and 60 nm $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ barriers repeated 20 times, all planar-doped with Si donors in the well centres at $2 \times 10^{10} \text{ cm}^{-2}$ and planar doped in the barrier centres at $3.5 \times 10^{10} \text{ cm}^{-2}$ (sample 1), $1.5 \times 10^{11} \text{ cm}^{-2}$ (sample 2), and $2.8 \times 10^{11} \text{ cm}^{-2}$ (sample 3).

Magnetotransmission spectra for sample 1 are shown in figure 2. The major features at 4.2 K (figure 2(a)) are CR

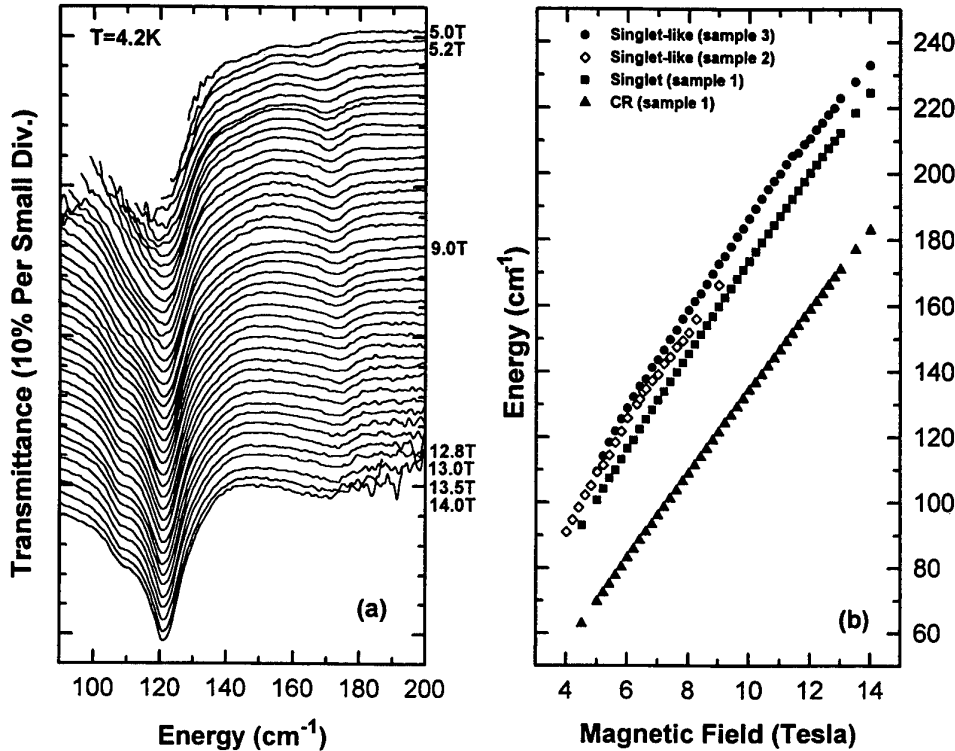


Figure 3. (a) Magnetotransmission spectra for sample 3 at 4.2 K in magnetic fields between 5 T and 13 T in steps of 0.2 T, and at 13.5 T and 14 T. (b) Transition-energies versus magnetic field for the D^- singlet (■) and CR (▲) for sample 1, and the singlet-like transitions for samples 2 (◇) and 3 (●).

at 121 cm^{-1} at 9 T (the CR lines for all the other spectra at different fields are aligned in order to show clearly the behaviour of the negative donor ion feature(s)). The strong feature at higher energies is the isolated D^- singlet line. The increasing separation from CR with increasing magnetic field (32 cm^{-1} to 42 cm^{-1}) is apparent for this sample (Holmes *et al* 1992). The positions of this transition versus magnetic field are plotted in figure 3(b) as the full squares. Figure 2(b) shows data taken on the same sample at 20 K. The strong feature that appears on the low-frequency side of CR is the lower-energy triplet transition (T_-) whose separation from CR is almost constant with magnetic field in this range ($12\text{--}13\text{ cm}^{-1}$ below CR); elevated temperature is required to populate the ground state of this transition (Ryu *et al* 1996). In addition, evidence of the $D^0|00\rangle\text{--}|10\rangle$ line is seen at higher frequencies. These data provide the reference positions for the isolated D^- singlet and triplet transitions for comparison with the many-electron bound magnetoplasmon transitions.

Magnetotransmission data for sample 3 (the sample with the highest excess electron density) are shown in figure 3(a). The singlet-like bound magnetoplasmon transition is clearly blue-shifted (approximately $9\text{--}10\text{ cm}^{-1}$ at 5 T) from the D^- singlet position of sample 1. In addition, the triplet-like bound magnetoplasmon is apparent at fields between 8 and 8.5 T as a weak shoulder on the low-field side of the broadened (mostly due to overabsorption) CR. Detailed line fitting for this sample and for sample 2 shows that this line is also slightly blue-shifted from its

isolated D^- triplet counterpart. However, since effects on the triplet-like bound magnetoplasmon are smaller and more difficult to extract due to its proximity to CR, the remainder of the discussion will concentrate on the singlet-like transition.

A summary of the resonance positions versus magnetic field for the D^- singlet and CR for sample 1, and the singlet-like magnetoplasmons for samples 2 and 3, is presented in figure 3(b). The substantial blue-shift of the singlet magnetoplasmon data of samples 2 and 3 above the D^- singlet data of sample 1 is apparent, with a larger shift for sample 3 (full circles) than for sample 2 (open diamonds). There are clear breaks in the slope for both samples 2 and 3, just below 6 T and around 11.5 T for sample 3, and just over 6 T for sample 2.

In order to examine these data more closely and to provide a direct comparison with theory, the data for samples 3 and 2 are replotted versus FF in figures 4(a) and (b), respectively, as energy differences between the singlet-like magnetoplasmon and the D^- singlet transitions normalized by the magnetic-field-dependent binding energy of a 2D neutral donor in the high-field limit, $E_0 = (\pi/2)^{1/2}(e^2/\epsilon\ell) = (\pi/2)^{1/2}E_{\text{Coul}}$, with ϵ the background dielectric constant and E_{Coul} the high-field Coulomb energy. The data for sample 3 at 4.2 K show clear cusp-like behaviour slightly above $\text{FF} = 1$ and slightly below $\text{FF} = 2$, in qualitative agreement with a strictly 2D, high-field-limit theory (Hawrylak 1994) which predicts a 2:1 ratio of the normalized difference at $\text{FF} = 2$ to that at $\text{FF} = 1$. The

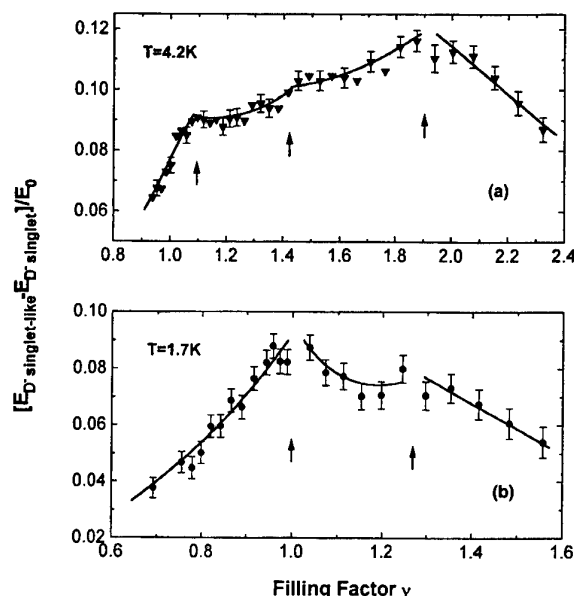


Figure 4. Energy differences between the singlet-like magnetoplasmon and the D^- singlet transitions normalized by the binding energy of a 2D neutral donor in the high-magnetic-field limit are plotted versus FF for samples 3 (a) and 2 (b). The lines are guides to the eye.

measured ratio is 1.3:1. The discrepancy is at least in part due to the finite well width (200 Å) and rather low fields of the experiments. The parameter that measures the field strength, $\gamma = \hbar\omega_c/2Ry^*$, where Ry^* is the effective Rydberg, varies from 0.6 to 2.2 over the experimental field range; so there is substantial Landau level mixing, which is not accounted for in the theory. There is also an indication of additional cusp-like behaviour near $FF = 1.4$, but this is not significantly outside experimental error.

The data for sample 2 at 1.7 K in figure 4(b) show even sharper cusp-like features near $FF = 1$, but do not encompass the low fields necessary to reach $FF = 2$ (< 2.9 T) for this lower-density sample. A comparison of the normalized shift at $FF = 1$ for this sample and sample 3 shows that the scaled blue-shift is identical for both samples (~ 0.09), which is indicative of the unifying aspects of the theoretical approach and clearly connects the

experimental observations to many-electron effects. The difference in behaviour between samples 2 and 3 in the region between $FF = 1$ and 1.5 may be related to the particularly low fields required for these FFs in sample 2 and overlap of the broadened Landau levels. Nevertheless, there is a clear anomaly in the vicinity of $FF = 4/3$ for this sample. At present there is no theoretical explanation for these anomalies, but from the behaviour near integral FFs, it is clear that these many-electron transitions are sensitive to the details of the electron (or quasi-electron) configuration, particularly in the vicinity of the positive charge. The present experiments clearly demonstrate that in the presence of a random array of positive donor ions many-electron effects are revealed directly in far-infrared magnetospectroscopy; these results should serve to motivate additional theoretical work in the region of fractional filling and additional work at lower temperatures and higher magnetic fields.

Acknowledgments

We are grateful to P Hawrylak for many illuminating discussions. The excellent MQW samples were grown by W Schaff at Cornell University. This work was supported in part by the Office of Naval Research under grants N00014-89-J-1673 and N00014-91-J-1939.

References

- Brey L, Fertig H A, Côte R and MacDonald A H 1995 *Phys. Rev. Lett.* **75** 2562
- Cheng J-P, Wang Y J, McCombe B D and Schaff W 1993 *Phys. Rev. Lett.* **70** 489
- Dzyubenko A B and Lozovik Yu E 1993 *Sov. Phys.-JETP* **77** 617
- Hawrylak P 1994 *Phys. Rev. Lett.* **72** 2943
- Holmes S, Cheng J-P, McCombe B D and Schaff W 1992 *Phys. Rev. Lett.* **69** 2571
- Huant S, Najda S P and Etienne B 1991 *Phys. Rev. Lett.* **65** 1486
- Johnson M H and Lippmann B A 1949 *Phys. Rev. B* **76** 828
- Kohn W 1961 *Phys. Rev.* **123** 1242
- Mandray A, Huant S and Etienne B 1992 *Europhys. Lett.* **20** 181
- Mueller E R, Larsen D M, Waldman J and Goodhue W D 1992 *Phys. Rev. Lett.* **68** 2204
- Ryu S R, Yu W-Y, Fu L P, Jiang Z X, Petrou A and McCombe B D 1996 *Proc. EP2DS-XI* (to be published in *Surf. Sci.*)
- Ryu S R, Jiang Z X, Li W J, McCombe B D and Schaff W 1996 *Phys. Rev. B* at press

Electrons in mesoscopically inhomogeneous magnetic fields

P D Ye[†], D Weiss^{†‡}, R R Gerhardt[†], G Lütjering[†],
K von Klitzing[†] and H Nickel[§]

[†] Max-Planck-Institut für Festkörperforschung, D-70569 Stuttgart, Germany

[‡] Experimentelle und Angewandte Physik, Universität Regensburg,
D-93040 Regensburg, Germany

[§] Forschungsinstitut der Deutschen Bundespost, D-64295 Darmstadt, Germany

Abstract. We investigated the low-field magnetoresistivity ρ_{xx} of a two-dimensional electron gas (2DEG) underneath microstructured ferromagnetic gratings. The strength and the direction of an externally applied magnetic field determines the strength and the shape of the micromagnet's stray field. Commensurability effects due to the imposed *periodic magnetic field* are in competition with similar effects resulting from a strain-induced *periodic electrostatic potential*, associated with the patterned ferromagnets on top of the heterojunction. The presence of both a periodic magnetic field and a periodic electrostatic potential gives rise to the interesting interference phenomena presented here.

By means of patterned ferromagnetic [1,2] or superconducting [3] layers on top of two-dimensional electron systems, it is possible to study electron motion in inhomogeneous magnetic fields varying on a length scale small compared to the mean free path l_e of the electrons. By periodically arranging ferromagnetic 'wires' with submicron diameters on top of a high-mobility GaAs–AlGaAs heterojunction (figures 1(a, c)), it is possible, for example, to generate a one-dimensional (1D) periodic magnetic field in the plane of the two-dimensional electron gas (2DEG). In analogy, two-dimensional (2D) periodic or random magnetic fields can be generated by placing ferromagnetic dots either periodically ([4], see also figures 1(b, d)) or randomly on top of a 2DEG [5]. Here, we focus on transport properties of a 2DEG underneath ferromagnetic gratings (figures 1(a, c)) resulting in both a 1D periodic magnetic field and a 1D periodic electrostatic potential. The latter is due to different thermal expansion coefficients of the ferromagnetic material (here dysprosium) and the semiconductor layers, which gives rise to a strain-induced modulation of the kinetic energy of the electrons [6].

The effect of a weak periodic magnetic field on a 2DEG has been predicted to result in an oscillatory magnetoresistance, which reflects the commensurability between the period a of the magnetic field modulation and the classical cyclotron diameter $2R_c$ of the electrons at the Fermi energy E_F [7–10]. For a weak 1D magnetic modulation with modulation amplitude $|B_m|$ much smaller than the external magnetic field $|B_0|$, the magnetoresistance ρ_{xx} oscillates with minima appearing at magnetic fields

given by [7–10]

$$2R_c = \frac{2\hbar k_F}{e|B_0|} = \left(\lambda + \frac{1}{4}\right)a \quad (1)$$

where $\lambda = 0, 1, \dots$ is an integer oscillation index, $k_F = \sqrt{2\pi n_s}$ is the Fermi wavenumber, n_s is the carrier density of the 2DEG, and a is the period of the 1D modulation in the x direction. This effect, observed recently [1–3], is intimately related to the commensurability (Weiss) oscillations observed in the resistivity ρ_{xx} of a 2DEG with weak electric modulation [11–13]. Similar to the electric case, the magnetic modulation modifies the energy spectrum [7–10]. The degenerate Landau levels are transformed into bands of finite width. The dispersion of these Landau bands provides an additional contribution to the resistivity ρ_{xx} , which vanishes only when the bandwidth becomes zero ('flat-band condition') [12, 13]. In contrast to equation (1), describing the flat-band condition for magnetic modulation, the flat-band condition for weak electrical modulation reads $2R_c = (\lambda - \frac{1}{4})a$ with $\lambda = 1, 2, \dots$. Hence, ρ_{xx} of a 2DEG with a weak electric modulation displays minima at B_0 fields where maxima would appear in a weak magnetic modulation of the same period a .

For the case of a pure electric modulation, the additional contribution to ρ_{xx} has been related to the classical guiding centre drift of the cyclotron orbits, which vanishes if the flat-band condition holds [14]. This classical picture can be extended to include, in addition to an arbitrary periodic modulation $V_m(\mathbf{r}) = \sum_{q \neq 0} V_q e^{iq\mathbf{r}}$ of the electrostatic potential energy of an electron, a weak modulation $B_m(\mathbf{r}) = \sum_{q \neq 0} B_q e^{iq\mathbf{r}}$ of the z component of the magnetic field ($q = (K n_x, K n_y)$) with $K = 2\pi/a$ and

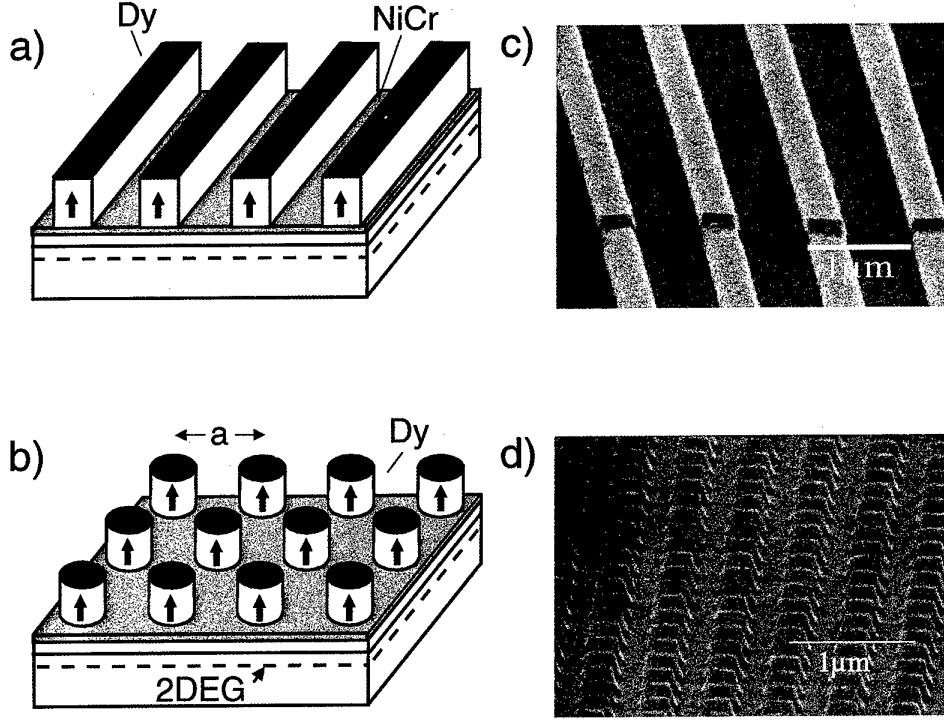


Figure 1. Sketch of ferromagnetic strips (a) and dots (b) used to generate a 1D and a 2D periodic magnetic field in the plane of a 2DEG in GaAs–AlGaAs heterojunctions. Arrows indicate a magnetization in the $\pm z$ direction. The corresponding electron micrographs show 200 nm high strips (c) and dots (d) made of dysprosium (Dy).

integers n_x, n_y , not simultaneously zero). Averaging the modulation-induced drift of the guiding centres over the unperturbed cyclotron orbits at field B_0 [14, 15], we obtain a change in the resistivity

$$\frac{\Delta\rho_{xx}}{\rho_0} = \frac{l_e^2}{2E_F^2} \sum_{q \neq 0} q_x^2 |S|^2 \quad (2)$$

$$S \approx \left(\frac{2}{\pi q R_c} \right)^{1/2} \left[\sigma V_q \cos \left(q R_c - \frac{\pi}{4} \right) + \frac{k_F}{q} \hbar \omega_q \sin \left(q R_c - \frac{\pi}{4} \right) \right] \quad (3)$$

where $\rho_0 = 1/en_s\mu$ is the zero-field resistivity of the unmodulated 2DEG with the electron mobility μ , $\omega_q = eB_q/m^*$ (m^* is the effective electron mass of GaAs) and σ is either +1 or −1 depending on the direction of the externally applied magnetic field B_0 . For our 1D modulation ($q = (Kn_x, 0)$), we find $\Delta\rho_{xy} = \Delta\rho_{yx} = \Delta\rho_{yy} = 0$ to leading order in the small parameter $(\mu B)^{-1}$.

Our samples were prepared from high-mobility GaAs–AlGaAs heterojunctions. The 2DEG was located approximately 100 nm underneath the sample surface. The carrier density n_s and electron mobility μ at 4.2 K were $\sim 2.2 \times 10^{11} \text{ cm}^{-2}$ and $1.3 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ in the dark, corresponding to an elastic mean free path of $\sim 10 \text{ } \mu\text{m}$. 50 μm wide Hall bars were fabricated by standard photolithographic techniques. Alloyed AuGe/Ni pads contact the 2DEG. A 10 nm thin NiCr film, evaporated on top of the devices, defines an equipotential plane to

avoid electric modulation of the 2DEG. However, strain due to different thermal expansion coefficients of the ferromagnetic grating and the heterojunction always results in a weak electric periodic potential as the sample is cooled down to cryogenic temperatures. The Dy gratings with periods of 500 nm and 1 μm were defined by electron beam lithography on top of the NiCr gates. After developing the exposed PMMA resist, a 200 nm Dy film was evaporated. After lift-off in acetone, ferromagnetic gratings, like the one shown in figure 1(c), were obtained. Four-point resistance measurements were performed in a ^4He cryostat with superconducting coils using standard a.c. lock-in techniques.

Figure 2 displays the magnetoresistance of a 2DEG measured for different directions of the magnetization M . The direction of M (see inset of figure 2) was adjusted by tilting the sample with respect to the direction of the external magnetic field which, in this case, was swept up to a maximum field of 10 T and then back to $B_0 = 0$ T. For the resistance measurements the samples were rotated back in such an orientation that the external magnetic field B_0 was applied normal to the plane of the 2DEG (z direction). Because the Dy microstructures used here show pronounced hard magnetic behaviour, the imposed magnetization is only slightly changed in the low-field region. If the strips are magnetized along their axis, essentially no magnetic modulation is imposed on the 2DEG, as expected, and the ρ_{xx} trace is virtually independent (on this scale) of the applied magnetic field. The situation changes dramatically if the strips are

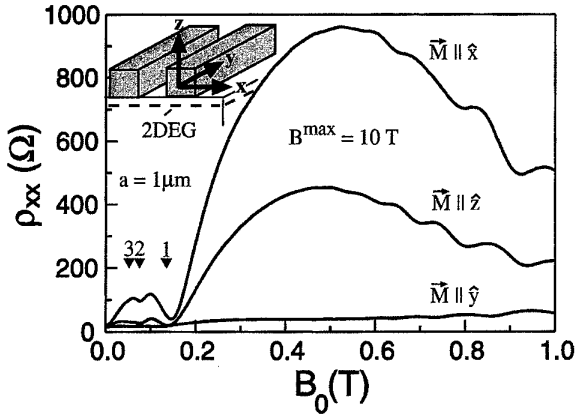


Figure 2. Resistance of a 2DEG as a function of the externally applied magnetic field B_0 (always pointing in the z direction) measured at 4.2 K. Different traces correspond to different directions of the magnetization M (see inset). The different amplitudes of the oscillations for $M \parallel \hat{x}$ and $M \parallel \hat{z}$ indicate different strengths of the stray field in the z direction and are probably a consequence of the rectangular shape of the wires (see figure 1(c)).

magnetized either in the z or in the x direction (see inset of figure 2). Pronounced oscillatory behaviour dominates ρ_{xx} with minima at magnetic field positions predicted for *pure* magnetic modulation. Hence, these traces display the *magnetic commensurability oscillations* anticipated by theory. Since on the other hand we know that a strain-induced electric modulation is present, we conclude that for the maximum ‘conditioning’ field of $B^{\max} = 10$ T the strength of our micromagnets covers up the effect of the electrostatic potential modulation.

By increasing the magnetization M gradually starting from $M = 0$, the influence of both the electrostatic periodic potential and the magnetic modulation can clearly be seen [1]. This is shown in figure 3 for a magnetization of the strips in the z direction where the strain-induced electrostatic periodic potential is *in phase* (maxima of the periodic potential and the periodic magnetic field are both underneath the centres of the magnetic strips; see figure 5(c)) with the periodic magnetic field. In figure 3 traces labelled (a) to (e) are taken for different maximum ‘conditioning’ fields B^{\max} between 1 T (a) and 10 T (e). With increasing B^{\max} a richer oscillatory structure unfolds in ρ_{xx} , indicating a growing amplitude of the magnetic field modulation. Open triangles mark the expected positions of the minima (flat-band condition) for pure electric modulation at $2R_c = (\lambda - \frac{1}{4})a$ while filled ones mark pure magnetic modulation at $2R_c = (\lambda + \frac{1}{4})a$. With increasing strength of the magnets (corresponding to growing B_q in equation (3)) the minima position shift from electric-modulation-dominated minima to magnetic-modulation-dominated minima (from 2_e to 1_m for $B_0 > 0$ and from 1_e to 1_m for $B_0 < 0$). The asymmetry of the traces with respect to $B_0 = 0$ is a consequence of changing the sign of σ in equation (3) [1]. The shift of the minima in figure 3 corresponds to a growing magnetic contribution $s_m \equiv (k_F/q)\hbar\omega_q \sin(qR_c - \frac{\pi}{4})$ in equation (3) whereas

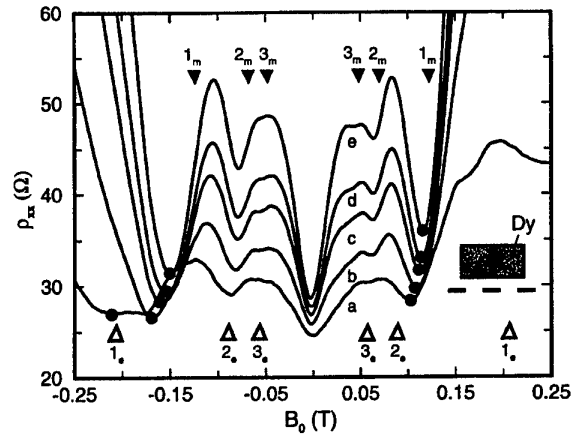


Figure 3. Low-field ρ_{xx} traces for a Dy grating with $a = 1 \mu\text{m}$ taken at 4.2 K for $M \parallel \hat{z}$ (see inset) showing the shift of the ρ_{xx} minima with increasing B^{\max} . Full triangles mark the position of the *magnetic* flat-band condition (subscript ‘m’) while the open triangles mark the *electric* ones (subscript ‘e’). The full circles highlight the positions of the ρ_{xx} minima.

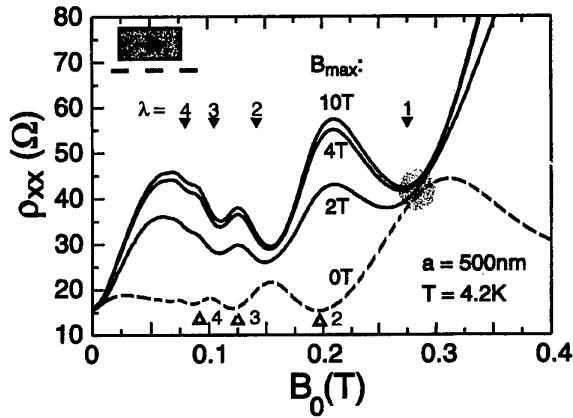


Figure 4. Low-field ρ_{xx} traces for a Dy grating with $a = 500 \text{ nm}$ measured at 4.2 K for $M \parallel \hat{x}$ (see inset). The dashed trace, taken after initial cooldown, reflects pure *electrostatic* commensurability oscillations. Solid traces are taken after magnetizing the wires (from $B^{\max} = 2$ T to 10 T) in the x direction. Contributions to ρ_{xx} resulting from increased magnetic modulation simply add to the electrostatically induced ρ_{xx} changes. This is clearly visible at the magnetic flat-band condition $\lambda = 1$ where the resistance minima sit upon the maximum of the dashed trace.

the strain-induced contribution, $s_e \equiv V_q \cos(qR_c - \frac{\pi}{4})$, is unaltered. The minima positions in ρ_{xx} are determined by the zeros of $|s_m + s_e|^2$. Hence, the shift can be used to estimate the amplitude of the magnetic stray field: assuming a simple sinusoidal modulation ($q = (2\pi/a, 0)$) we obtain a peak-to-peak modulation of 40 mT for trace (e) at $B_0 = 0$ [1].

The ‘interference’ of electric and magnetic modulation changes drastically if M is tilted in the x direction [15]. Now the magnetic modulation suffers a phase shift of $\pi/2$ with respect to the electric one (see figure 5(a)). Such a

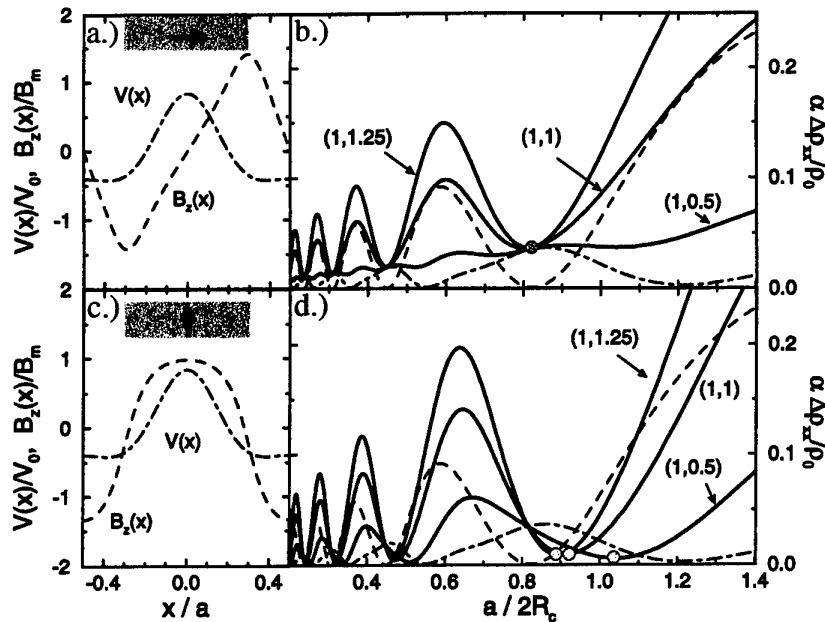


Figure 5. Left panel: normalized electric (dash-dotted) and magnetic (dashed) modulation fields calculated for a 1D lattice (period a) of strips of width $0.6a$ and height $0.4a$ magnetized either in the x (a) or in the z direction (c). The insets mark the positions and magnetization directions of the strips. The corresponding normalized and scaled (see [15]) magnetoresistance traces are shown in the right panels, where the dash-dotted curves refer to purely electric modulation, i.e. $(\epsilon, \mu) = (1, 0)$ and the dashed curves to purely magnetic modulation $(\epsilon, \mu) = (0, 1)$. The solid curves are calculated for a mixed modulation with the indicated (ϵ, μ) values.

phase shift, already addressed in [8], has consequences: while a magnetization in the z direction leads to real Fourier coefficients, that is to a pure cosine expansion of the modulation magnetic field, the magnetization in the x direction leads to purely imaginary coefficients, i.e. a sine expansion. As a result the Landau bands no longer become flat since now magnetic and electric modulation are simply added, $|s_m|^2 + |s_e|^2$. This additive behaviour can clearly be seen in the traces of figure 4 where the strips are magnetized parallel to the 2DEG. The dashed trace is taken directly after cooling down the device and the oscillations in ρ_{xx} are perfectly described by a pure electrostatic periodic potential, strain-imposed upon the 2DEG. With increasing B^{max} the effect of the magnetic field modulation takes over, as can be seen from the minima position at the magnetic flat-band condition (full triangles). At the magnetic flat-band condition $\lambda = 1$ near $B_0 \sim 0.28$ T it is obvious that the ρ_{xx} minimum 'sits' upon the maximum originating from the electric modulation.

Figure 5 displays calculations [15] illustrating the characteristic features observed in experiment. Figures 5(a) and 5(c) show the phase relation between the strain-induced potential $V_m(x)$ and the magnetic stray field $B_m(x)$. In figures 5(b) and 5(d) dashed and dash-dotted ρ_{xx} traces are calculated for pure magnetic and electric modulation, respectively. The solid curves are obtained for different normalized strengths μ of the magnetic modulation relative to a fixed electric modulation $\epsilon = 1$ (for a detailed description of the calculations and the definitions see [15]). For the in-phase situation in figure 5(d) the positions of the minima shift, as in experiment, when the amplitude of the

stray field is increased (from $\mu = 0.5$ to 1.25) from the electric flat-band condition at $a/2R_c \sim 1.3$ to the magnetic one at $a/2R_c \sim 0.8$ (corresponding to $B_0 < 0$ in figure 3). In contrast, for out-of-phase modulations the ρ_{xx} traces at magnetic flat-band conditions in figure 5(b) all meet in one point on top of the electric-modulation-induced ρ_{xx} maxima.

Acknowledgments

We gratefully acknowledge technical support by M Riek, B Schönherr and U Waizmann. The work was supported by the German Bundesministerium für Bildung und Forschung (BMBF).

References

- [1] Ye P D, Weiss D, Gerhardt R R, Seeger M, von Klitzing K, Eberl K and Nickel H 1995 *Phys. Rev. Lett.* **74** 3013
- [2] Izawa S, Katsumoto S, Endo A and Iye Y 1995 *J. Phys. Soc. Japan* **64** 706
- [3] Carmona H A, Geim A K, Nogaret A, Main P C, Foster T J, Henini M, Beaumont S P and Blamire M G 1995 *Phys. Rev. Lett.* **74** 3009
- [4] Ye P D, Weiss D, von Klitzing K, Eberl K and Nickel H 1995 *Appl. Phys. Lett.* **67** 1441
- [5] Ye P D *et al* unpublished. The influence of random magnetic fields on magnetotransport has been studied previously also by means of 'rough' macroscopic magnets:

- Mancoff F B, Clarke R M, Marcus C M, Zhang S C, Campman K and Gossard A C 1995 *Phys. Rev. B* **51** 13 269; or by using stochastically distributed vortices in a type-II superconducting film. See, e.g., Bending S J, von Klitzing K and Ploog K 1990 *Phys. Rev. Lett.* **65** 1060
- Geim A K, Bending S J, Grigorieva I V and Blamire M G 1994 *Phys. Rev. B* **49** 5749
- [6] Davies J H and Larkin I A 1994 *Phys. Rev. B* **49** 4800
- [7] Vasilopoulos P and Peeters F M 1990 *Superlatt. Microstruct.* **7** 393
- [8] Peeters F M and Vasilopoulos P 1993 *Phys. Rev. B* **47** 1466
- [9] Xue D P and Xiao G 1992 *Phys. Rev. B* **45** 5986
- [10] Yagi R and Iye Y 1993 *J. Phys. Soc. Japan* **62** 1279
- [11] Weiss D, von Klitzing K, Ploog K and Weimann G 1989 *Europhys. Lett.* **8** 179
See also *High Magnetic Fields in Semiconductor Physics II* (Springer Series in Solid State Sciences 87) ed G Landwehr (Berlin: Springer) p 357
- [12] Gerhardt R R, Weiss D and von Klitzing K 1989 *Phys. Rev. Lett.* **62** 1173
- [13] Winkler R W, Kotthaus J P and Ploog K 1989 *Phys. Rev. Lett.* **62** 1177
- [14] Beenakker C W J 1989 *Phys. Rev. Lett.* **62** 2020
- [15] Gerhardt R R 1996 *Phys. Rev. B* **53** at press

Universal conductance fluctuations in submicron wires of $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$

T Dietl[†], J Jaroszyński[†], G Grabecki[†], J Wróbel[†], M Sawicki[†],
T Skośkiewicz[†], E Kamińska[‡], A Piotrowska[‡], G Karczewski[†],
T Wojtowicz[†] and J Kossut[†]

[†] Institute of Physics, Polish Academy of Sciences, al. Lotników 32/46,
PL-02 668 Warszawa, Poland

[‡] Institute of Electron Technology, al. Lotników 32/46, PL-02 668 Warszawa, Poland

Abstract. Mesoscopic phenomena in nanostructures that incorporate diluted magnetic semiconductors may exhibit a number of novel features driven by spin interactions between mobile electrons and localized spins. Millikelvin studies of linear and nonlinear diffusive charge transport, which have been carried out for submicron wires of $n^+\text{-Cd}_{1-x}\text{Mn}_x\text{Te}$ epilayers as well as for microstructures of $\text{Hg}_{1-x-y}\text{Cd}_y\text{Mn}_x\text{Te}$ bicrystals, are reviewed. These studies have provided information on the significance of spin-disorder scattering and evidence of a new driving mechanism of the magnetoconductance fluctuations—the redistribution of the electrons between energy levels of the system, induced by the giant s–d exchange spin-splitting. Important implications of these findings for previous interpretations of spin effects in semiconductor and metal nanostructures are discussed.

1. Introduction

The most intriguing chain of discoveries over the last years has perhaps been that related to the quantum nature of electron diffusion in solids. The quantum effects are found to be responsible for the Anderson localization [1], unexpectedly large conductance fluctuations in small samples [2] and $1/f$ noise [3]. Furthermore, the accumulated findings have forced us to revise our intuition concerning Ohm's law, the Onsager relations and the role of energy dissipation and spin non-conserving processes. A wealth of theoretical expectations has been verified experimentally, mainly, in Si MOSFETs and GaAs/AlGaAs heterostructures [4].

In this paper we discuss briefly those aspects of quantum phenomena which can be of relevance in the case of diluted magnetic semiconductors (DMS). An interesting feature of DMS is a strong s–d exchange coupling between itinerant band electrons and electrons localized on magnetic shells. This coupling gives rise to spin-disorder scattering, giant spin-splitting of electronic states in magnetic fields, and the formation of magnetic polarons [5]. Furthermore, we present results of our experimental studies [6], which have recently been carried out on submicron wires of MBE grown $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$:In epilayers. Our findings, together with earlier experimental results for microstructures of narrow-gap $\text{Hg}_{1-x-y}\text{Cd}_y\text{Mn}_x\text{Te}$ bicrystals [7, 8], show that spin-disorder scattering gives rise to an additional phase-breaking mechanism but its destructive influence on the amplitude of universal conductance fluctuation (UCF) is much weaker than in metals. At the same time, the data

have led us to propose a new driving mechanism of the UCF in mesoscopic systems. This mechanism is associated with the giant spin-splitting and the corresponding redistribution of the electrons between energy levels of the system. The last part of the paper presents some implications of our work important in the context of previous interpretations of spin effects in semiconductor and metallic nanostructures.

2. Influence of magnetic impurities upon quantum transport

Soon after the discovery that the Anderson–Mott localization in disordered metals and the universal conductance fluctuations in mesoscopic conductors are controlled by diffusion poles in particle–particle and particle–hole correlation functions, it became clear [9–18] that the exchange coupling between the carriers and the subsystem of magnetic impurities, $-JsS$, should play an important role in the physics of quantum transport. Indeed, it is now well established that the localized spins, apart from introducing an additional temperature and magnetic-field-dependent contribution to the momentum relaxation rate, $1/\tau_s(T, H)$, can affect quantum transport phenomena in various ways, depending on their dynamics, ordering and relevant degrees of freedom [2, 3, 9–19]. We discuss separately the effects of spin-disorder scattering and spin-splitting.

2.1. Spin-disorder scattering

In a spin-glass phase, the perturbing potential associated with the frozen spins leads to a violation of the Onsager–

Büttiker symmetry relations in mesoscopic samples [13–15]. It also alters the scaling factor of the UCF amplitude [12, 16]. The fluctuating spins, on the other hand, because of an extreme sensitivity of the conductance to potential realizations, are expected to be an efficient source of the conductance noise [3, 11, 16, 17]. If, therefore, the integration time of the resistance meter is longer than the correlation time of the spin fluctuations and τ_s is shorter than the characteristic times of competing phase-breaking mechanisms, the exchange interaction results in a damping of the UCF amplitude [18–20]. Besides, electron scattering by disordered Ising spins gives rise to a cut-off in the particle–particle channels with $j_z = 0$ and the particle–hole channels with $j_z = \pm 1$, whereas the coupling to Heisenberg spins introduces a cut-off $1/\tau_s$ to all diffusion poles except for the particle–hole channel with total spin $j = 0$. This affects the quantum corrections to conductivity, and alters the universality class of the metal-to-insulator transition (MIT) [1].

Quantitatively, in the presence of a magnetic field, the relevant longitudinal and transverse spin relaxation rate of an electron liquid with a three-dimensional (3D) density of states assumes a well known form [5]:

$$\frac{1}{\tau_{s\parallel,\perp}} = \frac{k_B T \alpha^2 m^* k_F}{4\pi \hbar^3 g^2 \mu_B^2} \chi_{\parallel,\perp}(T, H) \quad (1)$$

where α is the s–d exchange integral; $\chi_{\parallel}(T, H) = \partial M(T, H)/\partial H$ and $\chi_{\perp}(T, H) = M(T, H)/H$. Here $M(T, H)$ is the magnetization of the Mn spins, which is usually well parametrized by a modified Brillouin function, $M(T, H) = \bar{x} N_0 g \mu_B S B_S(T + T_0, H)$, where $\bar{x} N_0$ and $T + T_0$ are the effective concentration and temperature of the localized spins respectively.

In the vicinity of the MIT, the spin–spin correlation function of the electron liquid becomes significantly renormalized by disorder and electron–electron interactions [21]. The apparent spin relaxation rate T_2 , as measured by the linewidth of electron spin resonance (ESR) or spin-flip Raman scattering (SFRS), is given by [22]

$$\frac{1}{T_2} = \left(\frac{2}{\tau_{s\parallel}} + \frac{2}{\tau_{s\perp}} \right) \left(1 + \frac{3\sqrt{3}}{\pi(k_F \ell)^2} \right) \left(1 + \frac{F}{2} \right) \quad (2)$$

where ℓ is the mean free path and F is the amplitude of Coulomb interaction [1].

2.2. Spin-splitting

The s–d coupling results in the spin-splitting of the electron states that depends on the temperature and the magnetic field according to [5]

$$\hbar\omega_s \equiv \tilde{g} \mu_B H = g^* \mu_B H + \alpha M(T, H)/g \mu_B. \quad (3)$$

The spin-splitting alters the universality class of the MIT [1] and the scaling factor of the UCF amplitude [12] as it introduces a cut-off in the particle–hole channels with $j_z = \pm 1$ and particle–particle channels with $j_z = 0$ [1, 23, 24]. This also leads to a giant positive magnetoresistance in the neighbourhood of the MIT [23, 25]. At the same

time, the spin-splitting leads to a redistribution of the carriers between the spin subbands. It was noted that the redistribution may affect the conductivity in two ways. First, by its influence on the mean free path that appears for $\hbar\omega_s \geq \varepsilon_F$ [26]. Second, by changing the distance of ε_F to the mobility edge ε_m [27], an effect operating if $\hbar\omega_s \geq |\varepsilon_F - \varepsilon_m|$.

Our findings, which we discuss below, show that in mesoscopic systems the redistribution of the carriers between the nanostructure energy levels constitutes a new driving mechanism of the UCF in a magnetic field. This is because the redistribution and the corresponding shift of the Fermi energy ε_F with respect to the bottom of the spin-down and spin-up subbands result in a gradual change of energy levels of the system which contribute to the conductance. The effect begins to show up at rather small values of the spin-splitting, $\hbar\omega_s \geq \max(k_B T, \hbar/T_2)$.

In order to evaluate the correlation field H_c of those magnetoconductance fluctuations that are induced by the spin-splitting, H_c^{spin} , we observe that, as long as $\varepsilon_F \gg \hbar\omega_s$, an increase of the magnetic field by ΔH leads to a shift of the Fermi energy $\Delta \varepsilon_F = \pm \frac{1}{2} \Delta H \partial \hbar\omega_s / \partial H$. Hence, for $\hbar\omega_s > \mu_c$, where [12] $\mu_c \approx \max(k_B T, \hbar/T_2)$ is the energy correlation range, we obtain H_c^{spin} in the form

$$H_c^{spin} = \sqrt{2} \mu_c (\partial \hbar\omega_s / \partial H)^{-1} \quad (4)$$

where the factor $1/\sqrt{2}$ appears since the fluctuations result from a superposition of two independent contributions associated with two different spin subbands. We presume that in magnetic materials the phase-breaking time $\tau_\varphi \equiv \hbar/\mu_c$ is equivalent to the spin-relaxation time T_2 of itinerant electrons. Since the correlation field for the orbital effects [2, 12] $H_c^{orb} \approx (\hbar/e)/(L_{min}^{(x)} \cdot L_{min}^{(y)})$, where $L_{min}^{(i)} = \min(L_\varphi, L^{(i)})$ and $L^{(x)} \cdot L^{(y)}$ is the sample area projected perpendicular to H , we see that the spin effect will dominate for sufficiently large values of the magnetic susceptibility $\chi(T, H) = \partial M(T, H)/\partial H$ and the inverse diffusion constant $D^{-1} = 3m^*/\hbar k_F \ell$.

3. Experimental results

3.1. $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$ epilayers

$\text{Cd}_x\text{Mn}_{1-x}\text{Te}$:In films with $x = 0$ or $x = 1 \pm 0.1\%$, a typical thickness of $0.3 \mu\text{m}$ and electron concentrations around 10^{18} cm^{-3} were grown by MBE on (001)-oriented semi-insulating GaAs epi-ready substrates with 10 \AA ZnTe and $3 \mu\text{m}$ CdTe undoped buffer layers. Secondary-ion mass spectroscopy (SIMS), high-resolution transmission electron microscopy (TEM), x-ray diffraction, photoluminescence, deep-level transient spectroscopy (DLTS), conductivity and Hall effect studies showed homogeneous impurity distribution and good structural properties of the epilayers [28]. In addition, SIMS and the room-temperature Hall data provided values of the Mn molar fraction and electron concentration, respectively, while their combination gave the activation of the In donors. The electron concentration in the studied samples is a factor of five greater than that corresponding to the MIT. The carrier and Mn

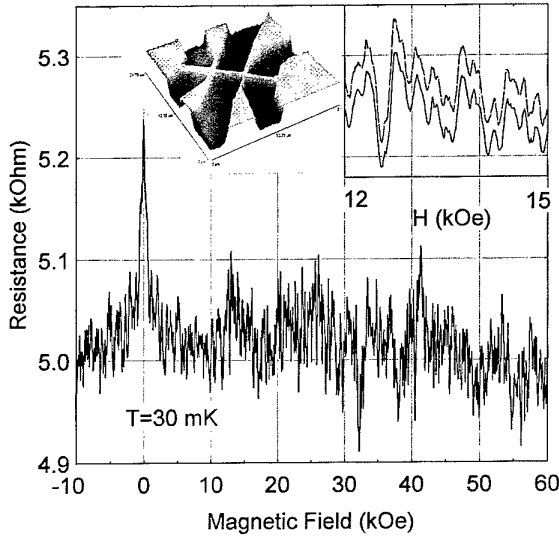


Figure 1. Resistance as a function of the magnetic field for the wire of n^+ -CdTe at 30 mK. The insets show an atomic force micrograph of the sample and the reproducibility of fluctuations (after [6]).

concentrations are low enough to preclude the appearance of either spin-glass freezing or the Kondo effect, even at the lowest temperature studied here, $T = 30$ mK [29].

The studied wires had the form of six-terminal Hall bars with an average width of $W \approx 0.3 \mu\text{m}$, and a distance between the voltage probes of $4\text{--}5 \mu\text{m}$, as shown in the insets to figure 1. They were fabricated by means of 30 keV electron-beam lithography, followed by wet etching in a 0.5% solution of Br_2 in ethylene glycol. Ohmic contacts were formed by alloying of indium. Low-frequency a.c. currents down to 100 pA were employed for the resistance measurements in a dilution refrigerator.

Figures 1 and 2 present the resistance as a function of the magnetic field perpendicular to wires of CdTe:In and $\text{Cd}_{0.99}\text{Mn}_{0.01}\text{Te}$:In with electron concentrations of 1.0×10^{18} and $8 \times 10^{17} \text{ cm}^{-3}$ respectively. Weak-field magnetoresistance and irregular reproducible resistance fluctuations are detected in both materials. We note that, because down to 100 mK the width W is bigger than the thermal diffusion length $L_T = \sqrt{\hbar D/k_B T}$, the studied wires are 3D with respect to phenomena that are sensitive to thermal broadening of the distribution function, such as electron–electron interactions. Since, however, in non-magnetic wires $L_\varphi = \sqrt{D\tau_\varphi} > L_T$ [23], we may expect a dimensional crossover in the negative magnetoresistance as it is controlled by phase-breaking effects. That this is indeed the case is shown in figure 2(a), which displays the temperature dependence of the magnetoresistance $\Delta\rho$ in n^+ -CdTe. The theoretical curves were calculated for the 3D case from the weak-localization theory [1], taking $m^*/m_0 = 0.099$ and assuming $L_\varphi = A/T^{3/4}$, where A was a fitting parameter determined to be $0.9 \mu\text{m K}^{3/4}$. Since in three dimensions $\Delta\rho \sim H^{-1/2}$, while in one dimension $\Delta\rho \sim H^{-1}$, we take the discrepancies between the experimental and calculated $\Delta\rho$, appearing at $L_\varphi(T) > W$, as evidence for the presence of the temperature-induced

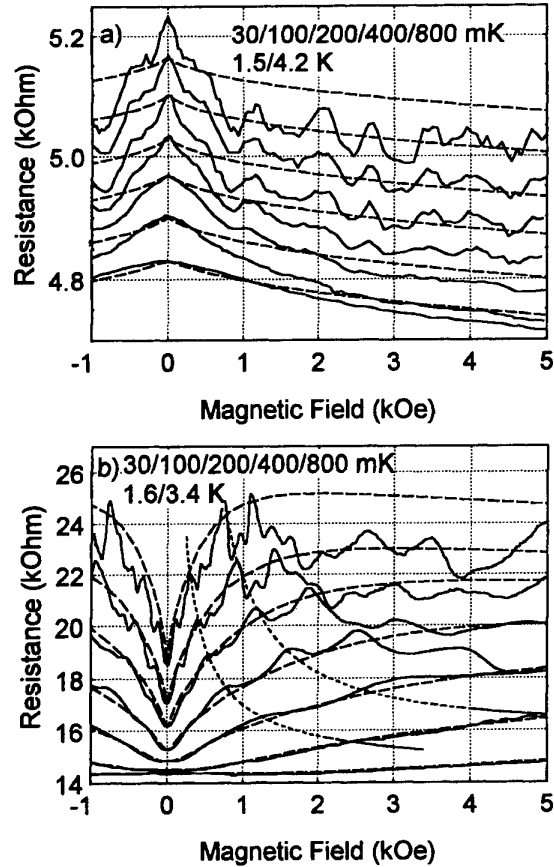


Figure 2. Resistance changes as a function of the magnetic field for the wires of n^+ - $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$ with $x = 0$ (a) and $x = 1\%$ (b) at various temperatures between 30 mK and 4.2 K (traces for the lowest temperatures are shifted upwards). Dashed curves represent the magnetoresistance calculated within the framework of 3D weak-localization theory [1, 23]. Dotted curves are guides for the eye, and visualize a strong temperature dependence of the resistance features in $\text{Cd}_{0.99}\text{Mn}_{0.01}\text{Te}$ (b) (after [6]).

crossover from three dimensions to one dimension at ~ 3 K in the studied wire.

A striking influence of the magnetic impurities upon the magnetoresistance and UCF is shown in figure 2(b), where data for n^+ - $\text{Cd}_{0.99}\text{Mn}_{0.01}\text{Te}$ are shown together with the results of a theoretical computation [1, 23]. As demonstrated previously [23], the positive magnetoresistance is caused by the effect of the giant exchange spin-splitting upon the electron–electron interaction. By taking well known material parameters of $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$ [5, 30], $\alpha N_0 = 0.22 \text{ eV}$, $N_0 = 1.48 \times 10^{22} \text{ cm}^{-3}$, $g = 2.0$, $S = \frac{5}{2}$ and, for $\bar{x} = 1\%$, $T_0 \approx 80 \text{ mK}$ [29, 30], we obtain $\hbar\omega_s(T, H)$, which gives $\bar{g}\mu_B H = 5.5 \text{ meV}$ for $Sg\mu_B H \gg k_B(T + T_0)$ and $\bar{g} = 150 \text{ K}/(T + T_0)$ in the opposite limit. The Coulomb amplitude was the only adjustable parameter and its fitting yielded $g_3 + g_4 = 1.3$. As expected, no dimensional crossover to one dimension is observed in the positive magnetoresistance as it is controlled by a short length scale, L_T .

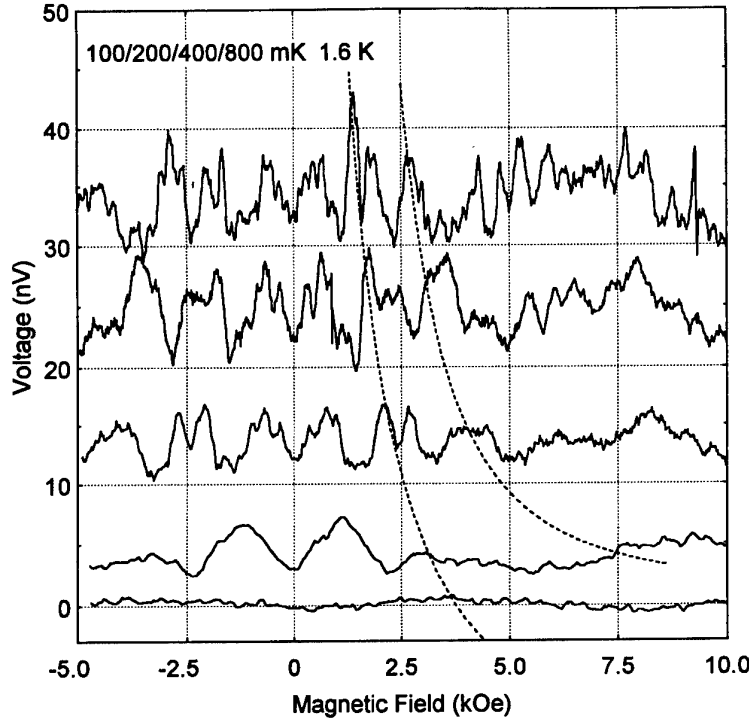


Figure 3. Voltage measured at a second harmonic of the excitation current (1 nA) as a function of the magnetic field. Broken curves are guides for the eye.

Turning to the resistance fluctuations in the studied samples we note that their root mean square (rms) amplitude is independent of the magnetic field. On the other hand, it increases with decreasing temperature according to $\text{rms}(\Delta R)/R^2 \approx (C/T)^r e^2/h$ where, above 100 mK, $C = 0.1$ mK and $r = 0.5$. While such behaviour is typical for non-magnetic one-dimensional wires, in which the distance between the voltage probes is greater than both L_T and L_ϕ [2, 12], it comes as a surprise in the case of $\text{Cd}_{0.99}\text{Mn}_{0.01}\text{Te}$. Indeed, in the latter, $\text{rms}(\Delta R)$ is expected to be controlled by τ_s , which for $T > T_0$ is independent of temperature but increases with the magnetic field [18–20]. However, because of the low density of states specific to semiconductors, we obtain $\hbar/k_B\tau_s = \alpha^2 m^* k_F \bar{x} N_0 S(S+1)/4\pi\hbar^2 k_B$ to be as low as 100 mK at $H = 0$. This estimate explains the minor importance of spin-disorder scattering in the studied system.

Another important aspect of the data depicted in figure 2 concerns the unusual behaviour of the correlation field H_c of the resistance fluctuations in $\text{Cd}_{0.99}\text{Mn}_{0.01}\text{Te}$. As shown by dotted curves, the fields H_i corresponding to characteristic points of the fluctuation pattern tend to increase with both temperature and magnetic field, a behaviour not observed in non-magnetic wires, including those of n^+ -CdTe. This new effect is visible not only for perpendicular but also for parallel orientation of the magnetic field with respect to the wires. Actually, we find that the dependence of H_i on H and T is best resolved in the fluctuations of the nonlinear response [2, 31], where the background magnetoresistance is absent. This is shown in figure 3, which displays fluctuation patterns of the voltage

taken at the second harmonics of the exciting a.c. current.

We note that field-induced changes of spin configurations have been proposed as the mechanism driving magnetoconductance fluctuations in spin-glass Cu:Mn wires [15]. We suggest instead that the dominant mechanism has its origin in the spin-splitting-induced redistribution of carriers between the spin subbands, an effect that can also operate in the paramagnetic phase considered here, as discussed in the previous section. Since, according to equation (3), the spin-splitting is proportional to the magnetization, the positions of the characteristic points of the fluctuation pattern should be temperature independent if the resistance is plotted as a function of the magnetization M , not of the magnetic field H . That this is indeed the case is shown in figure 4. Moreover, by putting parameters suitable for our CdMnTe wire ($k_F\ell = 1.5$ and $F = 2$), we obtain from equation (4), $H_c^{\text{spin}} = 410 \text{ T}[\text{K}]$ Oe. Except for the lowest temperatures, where the effect of a bound magnetic polaron may appear [22, 23], this agrees with the experimental values, $H_c^{\text{exp}} = 360 \text{ T}[\text{K}] + 36$ Oe, determined in the range $0.03 \text{ K} \leq T \leq 0.8 \text{ K}$ and $0 \leq H \leq k_B(T + T_0)/g\mu_B$.

3.2. $\text{Hg}_{1-x-y}\text{Cd}_y\text{Mn}_x\text{Te}$ bicrystals

It is now well established that defects associated with the grain-boundary plane in $p\text{-Hg}_{1-x-y}\text{Cd}_y\text{Mn}_x\text{Te}$ bicrystals have a donor character, and thus give rise to the formation of an inversion layer with a typical electron concentration of 10^{12} cm^{-2} . These electrons reside in a potential well of an average width of 150 \AA , occupy three to

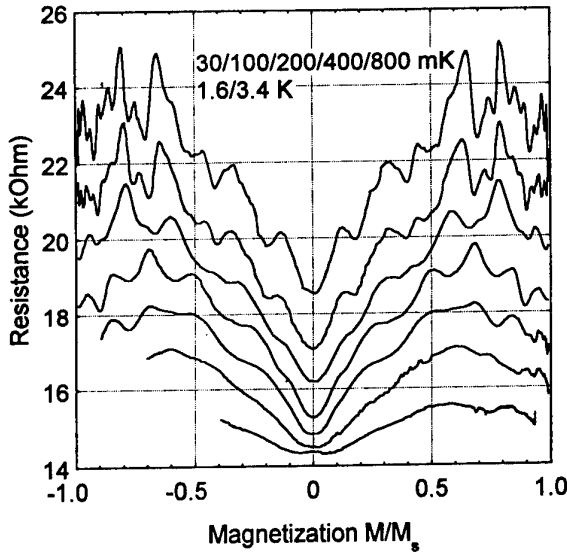


Figure 4. Magnetoresistance data of figure 2(b) plotted as a function of magnetization in units of $M_s = g\mu_B SN_0 \bar{x}$ (after [6]).

four electric subbands, and possess a mobility as large as $5 \times 10^4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, making the observation of the low-field Shubnikov-de Haas oscillations and the quantum Hall effect possible [32]. In order to examine properties of the macroscopic spin system via mesoscopic effects, this two-dimensional electron gas was probed by four microcontacts $\sim 15 \text{ }\mu\text{m}$ apart, an arrangement presented schematically in the inset to figure 5. For the studied Mn concentration, $x = 2\%$, according to our a.c. magnetic susceptibility measurements, the spin-glass freezing occurs at 100 mK, where inelastic and dephasing processes should be dominated by electron coupling to Mn spins.

As shown in figure 5, the studied samples exhibit a weak localization magnetoresistance superimposed on the universal conductance fluctuations [7]. For the small value of disorder involved ($k_F \ell \approx 50$), the positive magnetoresistance implies that the dominant spin-dependent scattering is associated with the spin-orbit coupling [1], as could be expected for narrow-gap semiconductors [33]. For the same reason (large $k_F \ell$ and small m^*) the UCFs are mainly driven by the orbital effect, not by the spin-splitting. On the other hand, spin-disorder scattering appears to be the main phase-breaking mechanism. It leads to a saturation of the magnitude of the magnetoresistance below 600 mK as well as to a saturation of the amplitude of the fluctuations below 200 mK, the temperature at which the phase-breaking length $L_\varphi \approx 2.5 \text{ }\mu\text{m}$ becomes shorter than the thermal length L_T . In the case of the magnetoresistance, the dephasing effect of Mn spins is associated with non-commutation of spin rotations corresponding to electron motion in the opposite directions along the same self-intersecting paths [10]. In contrast, the apparent amplitude of the fluctuations is diminished by Mn spins if the average time between their flips τ_f is shorter than the time constant of the resistance meter [18]. We conclude that for most Mn spins $\tau_f \leq 1 \text{ s}$, even in

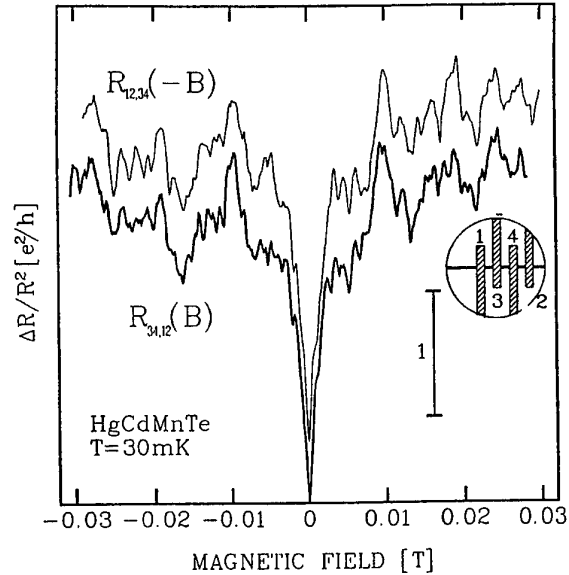


Figure 5. Resistance changes as a function of the magnetic field in a bicrystal of $\text{Hg}_{0.79}\text{Cd}_{0.19}\text{Mn}_{0.02}\text{Te}$ measured in a four-microprobe arrangement. Two traces were obtained by interchanging current and voltage probes and by reversing the direction of the magnetic field (after [7]).

the spin-glass phase, since we observed *no* increase of the fluctuation amplitude below T_f , the pattern was stable on a long time scale and the Onsager-Büttiker relations were fulfilled. In contrast, on cooling through T_f , an increase in the amplitude of the second harmonic generation has been noted [8]. The increase in deviations from Ohm's law (linear response) might be linked to the fact that scattering by frozen spins is essentially elastic, and thus below T_f the spin subsystem ceases to absorb energy from the carriers.

4. Discussion and conclusions

In summary, we have performed a low-temperature magnetoresistance study on structures for which the incorporation of magnetic impurities could be controlled during the growth process, while the details of magnetic properties and the value of the exchange coupling between the spins and the electrons could be monitored by means of a quantitative study of magneto-optical effects and weak-localization magnetoresistance. Our results indicate the minor importance of spin-disorder scattering in semiconductors. This, in particular, puts into question the suggestion [20] that spin-disorder scattering by etching-induced defects can account for the absence of resistance fluctuations in the region of weak magnetic fields in n^+ -GaAs wires at $T \geq 1.3 \text{ K}$. While we have not observed any signature of spin-glass freezing in linear transport, an increase of the nonlinear response upon cooling through T_f has been detected. This observation raises the question of the extent to which the magnitude of the noise [17] and the breakdown of the Onsager-Büttiker relations [14, 15] in Cu:Mn were affected by the deviations from Ohm's law.

Our data demonstrate the significant influence of the spin-splitting-induced redistribution of the carriers between the spin subbands on transport phenomena in mesoscopic systems. It would be interesting to find out whether the spin-splitting effect invoked here could account for the finite value of H_c^{spin} observed in Cu:Mn [15]. By taking parameters suitable for 1000 at.ppm Cu:Mn [15], i.e. $\ell = 200 \text{ \AA}$, $L_\varphi = 0.35 \text{ }\mu\text{m}$, $m^* = m_0$, $\chi = \chi_C(T_g) = 3.8 \times 10^{-4} \text{ emu}$ and $|\alpha N_0| = 0.5 \text{ eV}$, we obtain from equations (3) and (4) $H_c^{spin} = 4.1 \text{ kOe}$, a value quite close to the experimental finding, $H_c^{spin} = 4.2\text{--}6.4 \text{ kOe}$ [15].

Acknowledgments

The authors would like to thank R Duś and R Nowakowski for AFM visualization of our structures and Polish KBN for support under Contract No PBZ-Z011/P4/93/01.

References

- [1] Altshuler B L and Aronov A G 1985 *Electron-Electron Interaction in Disordered Systems* ed A L Efros and M Pollak (Amsterdam: North-Holland) p 1
- Fukuyama H 1985 *Electron-Electron Interaction in Disordered Systems* ed A L Efros and M Pollak (Amsterdam: North-Holland) p 155
- Lee P A and Ramakrishnan T V 1985 *Rev. Mod. Phys.* **57** 287
- Belitz D and Kirkpatrick T R 1994 *Rev. Mod. Phys.* **66** 261
- [2] Altshuler B L, Lee P A and Webb R A 1991 *Mesoscopic Phenomena in Solids* (Amsterdam: Elsevier)
- Washburn S and Webb R A 1992 *Rep. Prog. Phys.* **55** 1311
- [3] Feng S 1991 *Mesoscopic Phenomena in Solids* ed B L Altshuler, P A Lee and R A Webb (Amsterdam: Elsevier) p 107
- [4] Beenakker C V J and Van Houten H 1991 *Solid State Phys.* **44** 1
- [5] See, e.g., Dietl T 1994 *Handbook on Semiconductors* vol 3b, ed T S Moss (Amsterdam: North-Holland) p 1251
- [6] Jaroszyński J, Wróbel J, Sawicki M, Kamińska E, Skośkiewicz T, Karczewski G, Wojtowicz T, Piotrowska A, Kossut J and Dietl T 1995 *Phys. Rev. Lett.* **75** 3170; 1996 *Surf. Sci.* at press
- [7] Grabecki G, Jaroszyński J, Lenard A, Plesiewicz W, Skośkiewicz T, Dietl T, Kamińska E, Piotrowska A, Springholz G and Bauer G 1993 *Proc. 21st Int. Conf. on the Physics of Semiconductors (Beijing, 1992)* ed P Jiang and H-Z Zheng (Singapore: World Scientific) p 1407
- [8] Grabecki G, Dietl T, Plesiewicz W, Lenard A, Skośkiewicz T, Kamińska E and Piotrowska A 1994 *Physica B* **194-196** 1107
- [9] Lee P A 1980 *J. Non-Cryst. Solids* **35/36** 21
- Hikami S, Larkin A I and Nagaoka Y 1980 *Prog. Theor. Phys.* **63** 707
- [10] Bergmann G 1984 *Phys. Rep.* **107** 1
- [11] Altshuler B L and Spivak B Z 1985 *Pisma Zh. Eksp. Teor. Fiz.* **42** 363 (Engl. Transl. 1985 *JETP Lett.* **42** 477)
- Feng S, Bray A J, Lee P A and Moore M A 1987 *Phys. Rev. B* **36** 5624
- [12] Altshuler B L and Shklovskii B I 1986 *Zh. Eksp. Teor. Fiz.* **91** 220 (Engl. Transl. 1986 *Sov. Phys.-JETP* **64** 127)
- Lee P A, Stone A D and Fukuyama H 1987 *Phys. Rev. B* **35** 1039
- [13] Hershfield S 1991 *Phys. Rev. B* **44** 3320
- Liu J, Pennington M A and Giordano N 1992 *Phys. Rev. B* **45** 1267
- [14] de Vegvar P G N, Lévy L P and Fulton T A 1991 *Phys. Rev. Lett.* **66** 2380
- Weissman M B 1992 *Phys. Rev. Lett.* **68** 3468
- de Vegvar P G N and Lévy L P 1992 *Phys. Rev. Lett.* **68** 3485
- [15] de Vegvar P G N and Fulton T A 1993 *Phys. Rev. Lett.* **71** 3537
- [16] Cieplak M, Buřka B and Dietl T 1991 *Phys. Rev. B* **44** 12337; 1995 *Phys. Rev. B* **51** 8939
- [17] Israeloff N E, Weissman M B, Nieuwenhuys G J and Kasiorowska J 1989 *Phys. Rev. Lett.* **63** 794
- Israeloff N E, Alers G B and Weissman M B 1991 *Phys. Rev. B* **44** 12613
- [18] Bobkov A A, Falko V I and Khmel'nitskii D E 1990 *Zh. Eksp. Teor. Fiz.* **98** 703 (Engl. Transl. 1990 *Sov. Phys.-JETP* **71** 393)
- Chandrasekhar V, Santhanam P and Prober D E 1990 *Phys. Rev. B* **42** 6823
- Falko V I 1992 *J. Phys.: Condens. Matter* **4** 3943
- [19] Benoit A, Washburn S, Webb R A, Mailly D and Dumoulin L 1988 *Anderson Localization* ed T Ando and H Fukuyama (Berlin: Springer) p 346
- Van Haesendonck C, Vloeberghs H, Bruynseraede Y and Jonckheere R 1989 *Nanostructures: Physics and Fabrication* ed M A Reed and W P Kirk (Boston: Academic) p 467
- Benoit A, Mailly D, Perrier P and Nedellec P 1992 *Superlatt. Microstruct.* **11** 313
- Dietl T, Grabecki G and Jaroszyński J 1993 *Semicond. Sci. Technol.* **8** S141
- [20] Geim A K, Dubonos S V and Antonova I Yu 1990 *Pisma Zh. Eksp. Teor. Fiz.* **52** 873 (Engl. Transl. 1990 *JETP Lett.* **52** 247)
- [21] Sachdev S 1986 *Phys. Rev. B* **34** 6049
- [22] Dietl T, Sawicki M, Isaacs E D, Dahl M, Heiman D, Graf M J, Gubariev S I and Alov D L 1991 *Phys. Rev. B* **43** 3154
- [23] Sawicki M, Dietl T, Kossut J, Igalson J, Wojtowicz T and Plesiewicz W 1986 *Phys. Rev. Lett.* **56** 508
- [24] Ono Y and Kossut J 1984 *J. Phys. Soc. Japan* **53** 1128
- [25] Shapira Y, Oliveira N F Jr, Becla P and Vu T Q 1990 *Phys. Rev. B* **41** 5931
- [26] Shapira Y and Kautz R L 1974 *Phys. Rev. B* **10** 4781
- [27] Fukuyama H and Yosida Y 1981 *Physica* **105 B+C** 132
- Wojtowicz T, Dietl T, Sawicki M, Plesiewicz W and Jaroszyński J 1986 *Phys. Rev. Lett.* **56** 2419
- Aronzon B A, Chumakov N K, Dietl T and Wróbel J 1994 *Zh. Eksp. Teor. Fiz.* **105** 405 (Engl. Transl. 1994 *Sov. Phys.-JETP* **78** 216)
- [28] Karczewski G, Zakrzewski A, Kutrowski M, Jaroszyński J, Dobrowolski W, Grodzicka E, Janik E, Wojtowicz T and Kossut J 1995 *Acta Phys. Polon. A* **87** 241
- [29] Novak M A, Symko O G, Zheng D J and Oseroff S 1985 *J. Appl. Phys.* **57** 3418
- [30] Gaj J A, Grieshaber W, Bondin-Deshayes C, Cibert J, Feuillet G, Merle d'Aubigné Y and Wasiela A 1994 *Phys. Rev. B* **50** 5512
- [31] de Vegvar P G N, Timp G, Mankiewicz P M, Cunningham J E, Behringer R and Howard R E 1988 *Phys. Rev. B* **38** 4326
- de Vegvar P G N 1993 *Phys. Rev. Lett.* **70** 837
- [32] Grabecki G, Dietl T, Sobkowicz P, Kossut J and Zawadzki W 1984 *Appl. Phys. Lett.* **45** 1214
- Suski T, Wiśniewski P, Dmowski L, Grabecki G and Dietl T 1989 *J. Appl. Phys.* **65** 1203
- [33] Krenn H, Kaltenecker K, Dietl T, Spätek J and Bauer G 1989 *Phys. Rev. B* **39** 10918

Metal–insulator transition in Sb-doped short-period Si/SiGe superlattices

G Brunthaler[†], T Dietl[‡], M Sawicki[‡], G Stöger[†], J Jaroszynski[‡],
A Prinz[†], F Schäffler[†] and G Bauer[†]

[†] Institut für Halbleiterphysik, Johannes Kepler Universität, A-4040 Linz, Austria

[‡] Institute of Physics, Polish Academy of Sciences, al. Lotników 32/46, PL-02668 Warszawa, Poland

Abstract. Symmetrically strained Si/SiGe superlattices (25 Å wells/14 Å barriers) with homogeneous Sb doping concentrations of 4.5 and $6.0 \times 10^{18} \text{ cm}^{-3}$ show low-temperature magnetoconductivity effects which can be explained by single-electron backscattering and disorder-induced electron–electron interaction for an anisotropic three-dimensional case. The insulating phase of two lower doped samples (2.0 and $2.5 \times 10^{18} \text{ cm}^{-3}$) is documented by the observation of variable range hopping conductivity. In addition, the extrapolated zero-temperature conductivities for $4.5 \times 10^{18} \text{ cm}^{-3}$ in different magnetic fields show a magnetic-field-induced metal-to-insulator transition, demonstrating a three-dimensional behaviour of this superlattice with respect to localization.

1. Introduction

The metal-to-insulator transition (MIT) in semiconductors has been intensively investigated, as the disorder in the material can be easily tuned by the doping concentration. According to the Mott criterion $n_c^{1/3} a^* \approx 0.25$ (a^* is the effective Bohr radius), for doping concentrations below the critical value n_c , a semiconductor is in the insulating regime, i.e. its conductivity σ vanishes as the temperature T goes to zero ($\sigma\{T \rightarrow 0\} \rightarrow 0$). The investigations of Rosenbaum *et al* [1] in bulk Si:P have demonstrated that the doping-induced transition is continuous and that the quantum corrections [2–4] due to backscattering and disorder-modified electron–electron interaction are equally important.

By application of an external magnetic field, an MIT can also be induced in metallic samples. This effect is either due to a shrinkage of the wavefunctions or due to effects of spin splitting on the triplet states in the diffusion channel of the electron–electron interaction. The former effect is dominant in GaAs, for example, [5] as the small effective mass $m^* = 0.067m_0$ in this material leads to a large Bohr radius a^* and the small Landé factor $g^* = 0.4$ makes spin splitting less important. In Si, on the other hand, the ratio of cyclotron resonance to spin splitting m_0/m^*g^* , for $m^* = 0.19m_0$ and $g^* = 2.0$, is smaller by a factor 15 and makes the second effect dominant [6].

Apart from studies of the MIT in bulk semiconductors, a large amount of work has been done on low-dimensional systems. For the two-dimensional (2D) electron gas in Si inversion layers [7] the interplay between backscattering and electron–electron interaction was confirmed and evidence was found that all states are localized at low T .

Semiconductor superlattices (SLs), on the other hand, have been investigated to a much lesser extent. Superlattices possess several features different from bulk material.

(i) The band structure and thus the electronic properties are highly anisotropic.

(ii) The Fermi surface in reciprocal space is closed or open, depending whether the Fermi energy is smaller or larger than the miniband width w .

(iii) Fluctuations in layer thickness and barrier height provide an additional source for Anderson localization [8].

Most works concern experimental and theoretical studies of weak localization in GaAs/AlGaAs SLs [8, 9]. The magnetic-field-induced MIT in GaAs/AlGaAs was investigated recently by magnetoresistivity and infrared absorption measurements [10]. As a result of this study, the MIT occurs for B perpendicular to the layers at a lower magnetic field than for the parallel configuration. This is in perfect agreement with the shrinkage of the wavefunction due to its larger extent parallel to the layer planes.

2. Experiment

Low-temperature magnetotransport measurements were performed on four different n-type Si/SiGe SL samples with homogeneous Sb doping concentrations of $N_D = 2.0, 2.5, 4.5$ and $6.0 \times 10^{18} \text{ cm}^{-3}$ (sample identifications C1212, C1408, C1409 and C1410 respectively). The doping concentration was chosen in order to cover the range around the critical concentration of $2.9 \times 10^{18} \text{ cm}^{-3}$, at which the MIT in bulk Si:Sb was observed [11]. The Si/SiGe structures were grown by molecular beam epitaxy

on (001) substrates and consist of a 1 μm compositionally graded undoped $\text{Si}_{1-x}\text{Ge}_x$ buffer with x increasing from 0.05 to 0.22, a 0.5 μm undoped $\text{Si}_{0.83}\text{Ge}_{0.17}$ buffer, the homogeneously Sb-doped short-period SL with 150 periods (200 periods for C1212) of nominal 25 \AA Si wells and 15 \AA $\text{Si}_{0.55}\text{Ge}_{0.45}$ barriers and a 100 \AA cap layer. The SLs are strain symmetrized by the buffer, both having the same average Ge content. The concentration and homogeneity of the doping profile was checked by secondary-ion mass spectroscopy (SIMS) investigations. The structural properties of the SL were confirmed by the observations of the multilayer interference peaks in high-resolution x-ray diffraction; analysis of the spectra gives widths of 25 \AA and 14 \AA for wells and barriers respectively.

The magnetotransport investigations were performed on photolithographically patterned Hall bars in different cryostats (He^4 , He^3 and He^3/He^4 dilution refrigerator). In the He^4 system (magnetic field B up to 7 T) the sample could be rotated for investigating the angular dependence of the magnetoresistivity. In the He^3 system, fields up to 16 T can be achieved and in the dilution refrigerator lowest temperatures of 50 mK were reached at fields up to 8 T.

3. Results and discussion

Temperature-dependent resistivity and Hall effect measurements at 0.3 T from room temperature down to 10 K show a variation of only about 10% of the evaluated carrier concentration, whereas the mobility decreases by about a factor two and reaches values between 100 and 200 $\text{cm}^2 \text{V}^{-1} \text{s}^{-1}$ at 10 K. The estimated electron density n agrees to within 10% with the doping concentration from the SIMS analysis.

In the strain symmetrized Si/SiGe SLs, the biaxially strained Si layers form the conduction band (CB) wells in which the six-fold valley degeneracy is lifted. For the two lower-lying valleys the in-plane effective mass $m_{\parallel} = 0.19m_0$ is nearly a factor of 5 smaller than the perpendicular (longitudinal) mass $m_{\perp} = 0.92m_0$. The mobility for vertical transport is strongly reduced by the weak coupling of the SL quantum wells (QWs) through the 300 meV high $\text{Si}_{0.55}\text{Ge}_{0.45}$ barriers, resulting in a large anisotropy of electron transport.

According to our self-consistent subband calculations, in the absence of disorder only the lowest-lying subband is occupied. The ground state is located at $E_0 = 35$ meV above the CB edge in the Si wells, whereas the first excited state lies 100 meV higher in energy. (These findings were confirmed by intersubband absorption experiments.) The band structure is nearly independent of doping concentration as the absolute amount of charge transfer from the narrow doped barriers into the wells is small. The width of the ground state miniband $w = 2.5$ meV is less than the Fermi energy of $E_F = 11$ meV and hence the electron system possesses an open Fermi surface in reciprocal space. The in-plane electron mobility of about $\mu = 150 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$ at 10 K implies a momentum relaxation time, $\tau = m_{\parallel}\mu/e = 20$ fs, which is nearly identical to the vertical electron flight time, $\tau_f = d_w(m_{\perp}/2E_0)^{1/2} = 22$ fs, between the two interfaces in a single QW. Our SLs are in the limit of strong

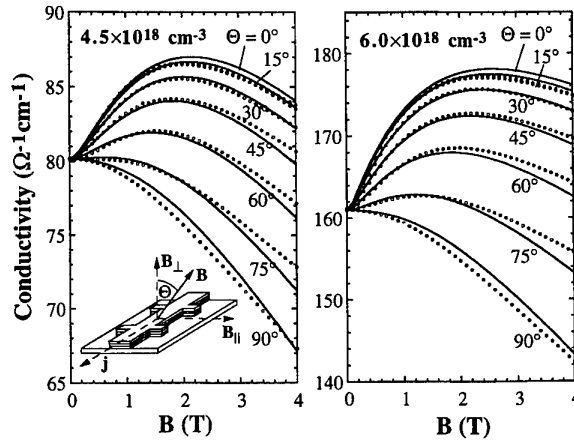


Figure 1. Angular dependence of the magnetoconductivity at 1.4 K for Si/SiGe superlattices with homogeneous Sb doping concentrations of 4.5 and $6.0 \times 10^{18} \text{ cm}^{-3}$. As indicated in the inset, 0° denotes the orientation of magnetic field perpendicular to the layers. The circles represent the experimental data, the full curves a least square fit according to equations (1)–(4) in the text.

disorder as the transport time τ corresponds to a scattering broadening of the density of states of $\hbar/\tau = 33$ meV, which is more than 10 times larger than the miniband width of $w = 2.5$ meV. This is thought to hinder vertical transport [12]. The product of Fermi wavevector and mean free path, $k_F l$, characterizes the disorder in the sense of diagrammatic perturbation theory [2–4]. For the sample with $4.5 \times 10^{18} \text{ cm}^{-3}$, we obtain $k_F l = 0.66$ at $T = 10$ K, which indicates a large distortion of the Fermi liquid picture as the value of $k_F l$ is less than 1.

Figure 1 shows the angular dependence of the magnetoconductivity for SLs with $n = 4.5$ and $6.0 \times 10^{18} \text{ cm}^{-3}$ at $T = 1.4$ K. The orientation of the magnetic field B is described by the angle Θ to the growth direction. The measurement current was perpendicular to the magnetic field for all angles. For B perpendicular to the layers, a positive magnetoconductivity at lower magnetic fields and a negative one at higher B is observed. For the parallel configuration only the negative component remains. The positive part is assigned to single-particle quantum interference from backscattering (orbital effect) whereas the negative is due to the Cooperon channel of the electron-electron interaction and the spin splitting of the triplet state in the diffusion channel [2–4].

The influence of the spin splitting on the magnetoresistance is not obvious but can be explained by the following arguments. The two-electron quasiparticles in the diffusion channel are formed via the disorder-induced electron-electron interaction. This quasiparticles can have either total momentum zero or one, the latter with components $-1, 0$ or $+1$ parallel to the magnetic field. Without magnetic field, all diffusions at the Fermi energy (only those contribute) have the same momentum k_F . As they have the same wavelength, all quasiparticles contribute to interference effects which create a contribution to the conductivity. For a magnetic field $g\mu_B B > k_B T$, the particles at the Fermi energy with different total momentum components

parallel to the magnetic field have different k_F and wavelength and thus a part of the interference effect is quenched which causes the negative magnetoresistance.

For intermediate fields, the strength of the 2D weak localization is proportional to the logarithm of the magnetic field, whereas for 3D it is proportional the square root of B . We found that in our case the 2D term failed to fit our data below 2 K, whereas the 3D version for anisotropic systems [2, 13, 14] is able to describe satisfactorily the field and angular dependence over the whole investigated magnetic field range.

The single-particle backscattering contribution (also called weak localization) is described by

$$\Delta\sigma_{xx}^{\text{WL}} = \alpha n_v \frac{e^2}{2\pi^2\hbar} \frac{D_{\parallel}}{D_a} \sqrt{\frac{eD_c B}{\hbar D_a}} f_3 \left\{ \frac{\hbar B}{4eD_c\tau_\phi} \right\} \quad (1)$$

with valley coupling parameter α , valley degeneracy $n_v = 2$, parallel diffusion constant $D_{\parallel} = E_F\mu/e$, the diffusion constant in the plane perpendicular to the magnetic field $D_c = [D_{\parallel}(D_{\perp}\sin^2\Theta + D_{\parallel}\cos^2\Theta)]^{1/2}$, the averaged diffusion constant $D_a = (D_{\parallel}^2 D_{\perp})^{1/3}$, the phase coherence time τ_ϕ and the functional dependence f_3 according to equation (7.1) in Altshuler and Aronov [2]. In D_c and D_a the perpendicular diffusion constant D_{\perp} also enters, which in our case describes the motion of electrons between neighbouring QWs.

The spin splitting of the diffusion term is given by

$$\Delta\sigma_{xx}^{\text{D}} = -g_3 n_v^2 \frac{e^2}{8\pi^2\hbar} \sqrt{\frac{k_B T}{2\hbar D_{\perp}}} f_D \left\{ \frac{g^* \mu_B B}{k_B T} \right\} \quad (2)$$

with g_3 the diffusion amplitude [3] and f_D given by an integral representation according to equation (6.47) in [2].

In addition to the single-electron backscattering and the diffusion (particle-hole) channel, the Cooperon (particle-particle) channel was taken into account by

$$\Delta\sigma_{xx}^{\text{C}} = -g_4 \frac{e^2}{4\pi^2\hbar} \sqrt{\frac{k_B T}{\hbar D_{\perp}}} f_C\{x1, x2, x3\} \quad (3)$$

with the arguments $x1 = 2\hbar D_c/\pi l_B^2 k_B T$, $x2 = g^* \mu_B B/\pi k_B T$ and $x3 = \hbar/\pi\tau_\phi k_B T$ of the Cooperon function f_C (equation (6.42) in [2]) and the Coulomb amplitude [3, 15]

$$g_4 = \frac{g_3}{1 + g_3 \ln(E_F/\varepsilon)} \quad (4)$$

with $\varepsilon = \max\{k_B T, 2\hbar D_c/\pi l_B^2, g^* \mu_B B/\pi\}$ and $l_B = (\hbar/eB)^{1/2}$ the magnetic length. Here it was assumed that intervalley scattering reduces the valley degeneracy factor αn_v to 1 as observed in analysis of the single-particle interference term in Si/SiGe QWs with mobilities in the range of $10^4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ [16].

The electron Landé factor g^* is expected to be isotropic and should not be rescaled due to correlation effects in confined electron systems [2]. This is in contrast to the spin splitting which enters into the magnetic susceptibility, where a renormalized value of g^* should be used [17]. For our analysis, we take the bulk value of n-Si, $g^* = 2.0$.

The fit of the magnetic field and angular dependence, as indicated by the full curves in figure 1, gives for $n = 4.5 \times 10^{18} \text{ cm}^{-3}$ (and $6.0 \times 10^{18} \text{ cm}^{-3}$ respectively) the anisotropy factor $r_D = D_{\parallel}/D_{\perp} = 40(108)$, phase coherence time $\tau_\phi = 13.3 \text{ ps}$ (14.6 ps) and the spin-diffusion factor $g_3 = 1.1$ (1.3). All other parameters in equations (1)–(4) follow directly from these fitted values and from electron concentration n and temperature T .

As can be seen from the good quality of the fit in figure 1, the standard 3D theory of weak localization is able to describe the behaviour of our samples with higher doping concentrations. The magnitude of the magnetoconductance is proportional to the square root of the anisotropy factor r_D and is thus larger than in bulk 3D systems. This interesting property has already been noted in the context of lateral transport in GaAs/AlGaAs SLs [9].

From our data, the vertical diffusion constant $D_{\perp} = D_{\parallel}/r_D$ can be deduced, which is difficult to be measured directly in Si/SiGe SLs due to problems with doping of the graded buffer layer. From D_{\perp} the escape time from one well into another can be estimated. For $\hbar/\tau > w$, the subsequent transitions between wells are uncorrelated, so that $\tau_{\text{esc}} \approx (d_w + d_b)^2/D_{\perp}$. This gives $\tau_{\text{esc}} = 5.1 \text{ ps}$ and 6.9 ps for samples with 4.5 and $6.0 \times 10^{18} \text{ cm}^{-3}$ respectively, a value much longer than the escape time for coherent coupling between the wells, $2\hbar/w = 0.52 \text{ ps}$. Due to the strong disorder $\hbar/\tau > w$, the coherent transport is hindered and the transition rate should be described by a tunnelling probability through a barrier with width d_w . This gives an estimate of the escape time of 9 ps , assuming that only the confinement energy E_0 determines the tunnelling energy and that the parallel momentum is conserved. The escape time evaluated from D_{\perp} is lower, which may come from additional momentum non-conserving tunnelling processes [18].

The magnetic field and angular dependence of the magnetoconductivity were also investigated as a function of temperature up to 20 K. Holding the rate r_D and the amplitude g_3 constant during the least square fit of the data, we get a temperature dependence of $\tau_\phi \propto T^{-\beta}$ with $\beta = 1.0$ and 1.2 for $n = 4.5$ and $6.0 \times 10^{18} \text{ cm}^{-3}$ respectively. With increasing temperature, the deviation between the calculated and measured σ increases. When we allow also a variation of the parameter r_D during the fit for higher T , a good agreement can again be achieved. In that case the fitted values of $\tau_\phi(T)$ remain practically unchanged but the rate r_D decreases from 40 and 108 for the two SL systems at 1.4 K down to 26 and 37 respectively, at 20 K. As D_{\parallel} increases only slightly with T , this corresponds to an increase of D_{\perp} and thus to a higher transition rate between QWs for higher T . This behaviour cannot be understood with momentum-conserving tunnelling processes alone, but might be possible for disorder-assisted inelastic tunnelling.

Only the SL systems with higher electron concentration (4.5 and $6.0 \times 10^{18} \text{ cm}^{-3}$) could be well described by equations (1)–(4) at $T \geq 1.4 \text{ K}$. For the lower-doped samples (2.0 and $2.5 \times 10^{18} \text{ cm}^{-3}$) the quantum corrections become very large and the assumptions of small perturbation and $k_F l \geq 1$ are no longer valid. This is demonstrated in figure 2 for $n = 2.0 \times 10^{18} \text{ cm}^{-3}$. The

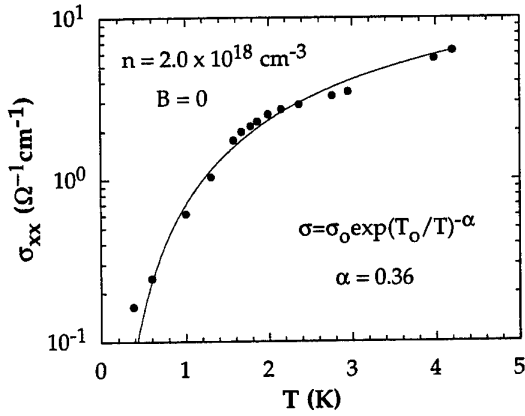


Figure 2. The SL structure with the lower doping concentration of $2.0 \times 10^{18} \text{ cm}^{-3}$ shows a strong decrease of conductivity towards lower temperatures, which can be described by $\sigma = \sigma_0 \exp[-(T_0/T)^\alpha]$ typical for variable-range hopping in the insulating phase.

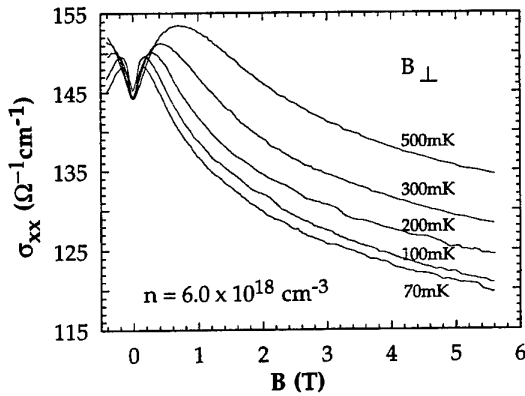


Figure 3. Magnetoconductivity of an Si/SiGe superlattice with Sb doping concentration of $6.0 \times 10^{18} \text{ cm}^{-3}$ at temperatures between 70 mK and 500 mK.

conductivity at zero magnetic field decreases dramatically at lower temperatures and can be described by $\sigma = \sigma_0 \exp[-(T_0/T)^\alpha]$ with $\alpha = 0.36$ and $T_0 = 100 \text{ K}$. This behaviour is typical of a variable-range hopping mechanism and demonstrates that the lower-doped samples are in the insulating phase.

In contrast, the sample with $n = 6.0 \times 10^{18} \text{ cm}^{-3}$ exhibits metallic conductivity as shown in figure 3. The zero magnetic conductivity does not change at all down to 70 mK, and the magnetoconductivity up to 6 T decreases only moderately, giving no indication of a magnetic-field-induced MIT either.

As will be demonstrated in the following, the SL with an electron concentration of $4.5 \times 10^{18} \text{ cm}^{-3}$ is the most interesting one. Its magnetoconductivity is shown in figure 4 versus the square root of T in the millikelvin temperature region for (a) magnetic field perpendicular to the SL layers and (b) parallel. A dependence of σ on $T^{1/2}$ is observed over a wide temperature range, characteristic of 3D systems [1, 11, 19, 20]. For the perpendicular

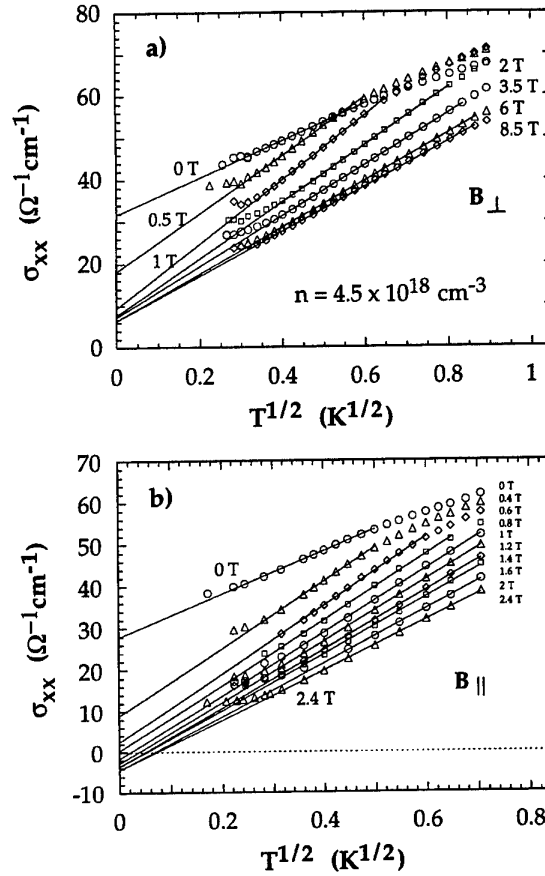


Figure 4. Temperature dependence of the conductivity for the sample with $4.5 \times 10^{18} \text{ cm}^{-3}$ at different magnetic fields on a square root of T scale for the (a) perpendicular and (b) parallel orientation of the magnetic field.

configuration (figure 4(a)) the conductivity extrapolated to $T = 0 \text{ K}$ is clearly above zero for all investigated fields and even seems to saturate. On the other hand, the parallel configuration (figure 4(b)) shows extrapolated conductivities below zero, indicating a magnetic-field-induced MIT transition. The transition occurs for magnetic fields between 1 and 2 T. The conductivity data show a deviation from the straight lines at the lowest temperatures. This may come from a thermal decoupling of the electron gas from the cold source; similar effects were observed in recent investigations of bulk Si:P [19]. On the other hand a similar anomaly was observed in bulk Si:Sb and assigned to a possible electronic phase transition or a spontaneous lifting of the valley degeneracy from 6 to 2 [11]. The latter explanation can be excluded by our experiment as the lifting of the valley degeneracy already occurs due to the biaxial strain in the Si/SiGe SLs.

As obtained by the fitting of the magnetoconductivity for the two higher-doped structures (figure 1), the decrease in conductivity in the parallel configuration is due to the diffusion-spin term whereas for the perpendicular orientation this decrease is partly counterbalanced by an increase due to the destruction of the single-electron backscattering effect. Thus the observation of the field-

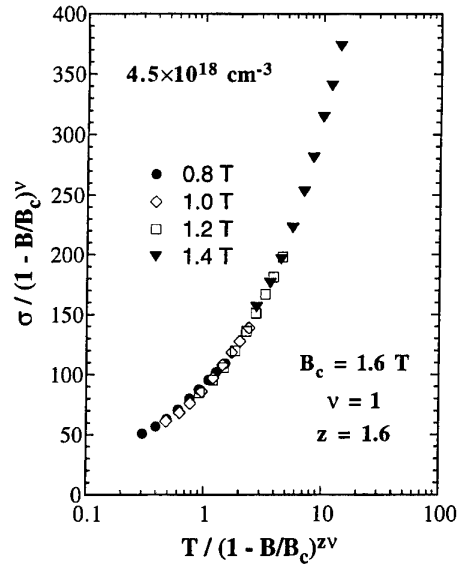


Figure 5. Scaling behaviour of the magnetoconductivity for $4.5 \times 10^{18} \text{ cm}^{-3}$ in the vicinity of the metal-to-insulator transition for $B = 0.8, 1.0, 1.2$ and 1.4 T according to $\sigma / (1 - B/B_c)^\nu = T / (1 - B/B_c)^{z\nu}$.

induced MIT in the parallel configuration proves rather directly that in Si/SiGe structures the transition is driven by the spin term and not by orbital effects. This is in contrast to n-GaAs with a large ratio of cyclotron to spin-Zeeman energy $m_e/m^*g^* = 37$ (for Si 2.6). Indeed, in n-GaAs/GaAlAs SLs the MIT occurs at lower magnetic field when it is oriented perpendicular to the layers [10] indicating the dominance of orbital effects (wavefunction shrinkage).

In order to find out whether the behaviour of the conductance in the vicinity of the MIT is consistent with theoretical predictions for 3D systems, we replot the experimental results taken between 90 and 500 mK in the coordinates suggested by Finkelstein's renormalization group approach [21, 22]

$$\frac{\sigma(B, T)}{(1 - B/B_c)^\nu} = F\left(\frac{T}{(1 - B/B_c)^{z\nu}}\right) \quad (5)$$

where B_c is the critical magnetic field, ν the correlation length and z the dynamic exponent. As shown in figure 5, the experimental results agree well with the scaling law, but it turns out that due to the strong dependence of σ versus T our data are not sensitive to the choice of $B_c \approx 1.6 \text{ T}$ and $\nu \approx 1$ but provide an accurate value of $z = 1.6 \pm 0.2$. A similar analysis for bulk n-Si at $B = 0$ yielded $z = 3.1$ [23], while that for the MIT of spin-polarized electrons in diluted magnetic semiconductors (DMS) results in $z = 1.7 \pm 0.4$ [24]. This is consistent, as DMS and n-Si in the magnetic field belong to the same universality class, whereas n-Si without a magnetic field belongs to another one [21, 22, 24].

4. Conclusions

The magnetic field and angular dependence of the conductivity of Si/Si_{0.55}Ge_{0.45} SL structures with electron concentrations of 4.5 and $6.0 \times 10^{18} \text{ cm}^{-3}$ can be well described by contributions from backscattering and disorder-induced electron-electron interaction effects for anisotropic three-dimensional systems using a rescaled diffusion constant $D_a = (D_\parallel^2 D_\perp)^{1/3}$. For the orientation of the magnetic field parallel to the SL layers, the spin splitting of the diffusion-triplet state dominates the magnetoconductivity. For the samples with lower electron concentrations (2.0 and $2.5 \times 10^{18} \text{ cm}^{-3}$) the insulating behaviour at low temperatures is demonstrated by the observed conductivity in the variable-range hopping regime. The magnetoconductivity of the higher-doped samples in the millikelvin regime follows a $T^{1/2}$ law over a wide range. This, together with the observed MIT, implies a three-dimensional behaviour of the studied SL systems.

We suggest that at non-zero temperature an SL remains three dimensional as long as an electron can move from one well to another before the phase coherence is destroyed, i.e. the escape time τ_{esc} must be shorter than both the phase coherence time τ_ϕ and the thermal coherence time $\hbar/k_B T$. In our SLs, this condition is fulfilled in the investigated temperature range below 2 K .

Acknowledgments

We would like to thank B L Altshuler, M Helm and P Wölflé for helpful discussions. This work was supported by FWF Vienna, Austria, Polish KBN and Austrian Ost-West Programm.

References

- [1] Rosenbaum T F, Milligan R F, Paalanen M A, Thomas G A, Bhatt R N and Lin W 1983 *Phys. Rev. B* **27** 7509
- [2] Altshuler B L and Aronov A G 1985 *Electron-Electron Interactions in Disordered Systems* ed A L Efros and M Pollak (Amsterdam: North-Holland) p 1
- [3] Fukuyama H 1985 *Electron-Electron Interactions in Disordered Systems* ed A L Efros and M Pollak (Amsterdam: North-Holland) p 155
- [4] Lee P A and Ramakrishnan T V 1985 *Rev. Mod. Phys.* **57** 287
- [5] Maliepaard M C, Pepper M, Newbury R and Hill G 1988 *Phys. Rev. Lett.* **61** 369
- [6] Dai P, Zhang Y, Bogdanovich S and Sarachik M P 1993 *Phys. Rev. B* **48** 4941
- [7] Uren M J, Davies R A and Pepper M 1980 *J. Phys. C: Solid State Phys.* **13** L985
- [8] Lee M, Solin S A and Hines D R 1993 *Phys. Rev. B* **48** 11 921
- [9] Szott W, Jedrzejek C and Kirk W P 1993 *Phys. Rev. B* **48** 8963 and references therein
- [10] Hilber W, Helm M, Peeters F M, Alavi K and Pathak R N 1996 *Phys. Rev. B* **53** 6919
- [11] Long A P and Pepper M 1984 *J. Phys. C: Solid State Phys.* **17** L425
- [12] Finlayson D M, Mason P J and Tunstall D P 1990 *J. Phys.: Condens. Matter* **2** 6735
- [12] See, for example, Yang S-R E and Das Sarma S 1988 *Phys. Rev. B* **37** 10 090

- [13] Altshuler B L, Aronov A G, Larkin A I and Khmel'nitskii D E 1981 *Zh. Eksp. Teor. Fiz.* **81** 768 (1981 *Sov. Phys.-JETP* **54** 411
See also Ootuka Y and Kawabata A 1985 *Prog. Theor. Phys. (Kyoto)* suppl **84** 249
- [14] Bhatt R N, Wölfe P and Ramakrishnan T V 1985 *Phys. Rev. B* **32** 569
- [15] Sawicki M, Dietl T, Kossut J, Wojtowicz T and Plesiewicz W 1986 *Phys. Rev. Lett.* **56** 508
- [16] Brunthaler G to be published
- [17] Isihara A 1989 *Solid State Phys.* **42** 271
- [18] Movaghar B, Leo J and MacKinnon A 1988 *Semicond. Sci. Technol.* **3** 397
- [19] Stupp H, Hornung M, Lakner M, Madel O and von Löhneysen H 1993 *Phys. Rev. Lett.* **71** 2634
- Hornung M, Ruzzu A, Schlager H G, Stupp H and von Löhneysen H 1994 *Europhys. Lett.* **28** 43
- [20] Dai P, Zhang Y, Bogdanovich S and Sarachik M 1993 *Phys. Rev. B* **48** 4941
- [21] See Belitz D and Kirkpatrick T R 1994 *Rev. Mod. Phys.* **66** 261
- [22] Finkelstein A M 1984 *Zh. Eksp. Teor. Fiz.* **86** 367 (1984 *Sov. Phys.-JETP* **59** 212)
Castellani C, Di Castro C, Lee P A and Ma M 1984 *Phys. Rev. B* **30** 527
- [23] Stupp H, Hornung M, Lakner M, Madel O and von Löhneysen H 1994 *Phys. Rev. Lett.* **72** 2122
- [24] Wojtowicz T, Dietl T, Sawicki M, Plesiewicz W and Jaroszynski J 1986 *Phys. Rev. Lett.* **56** 2419 (the critical exponent z is related to G^* determined in that paper by $z = 3 - 2/G^*$)

A study of the interacting electron system in stressed (001) Si MOSFETs

J Lutz and F Kuchar

Institute of Physics, University of Leoben, Franz Josef Straße 18, A-8700 Leoben, Austria

Abstract. The interaction of the electrons in Si MOSFETs has been studied under uniaxial stress using quantum magnetotransport experiments (Shubnikov–de Haas and quantum Hall effect). The stress allows one to vary the relative positions of the conduction band valleys where the electron–electron interaction plays a crucial role. From the evaluation of the data we obtain the phase diagram of the population of the valleys as a function of stress and carrier density. Our results are excellently described by the theory of Takada and Ando where intervalley electron–electron interaction is taken into account.

1. Introduction

The study of electron–electron interaction effects is one of the interesting topics of solid state physics. In silicon the electrons are accommodated in six equivalent valleys of the conduction band. Therefore, intervalley interaction has to be considered. The energetic degeneracy of the valleys is lifted by a confinement potential leading to two sets of valleys in the two dimensional electron system (2DES) with circular and elliptic constant energy contours respectively (figure 1).

The intervalley interaction can be studied in this system by the application of uniaxial stress that shifts the valleys against each other and this allows the interaction to be tuned. The individual shifts of the valleys relative to a common hydrostatic term are given by [1, 2]

$$\Delta E_{\alpha} = \pm D_u X (S_{11} - S_{12}) [(\hat{k}_{\alpha} \cdot \hat{F})^2 - \frac{1}{3}].$$

The + and – signs denote tensile and compressive stress respectively. \hat{F} is the unit vector in the direction of the applied force which produces the stress of value X . For silicon $D_u = 9$ eV and $S_{11} - S_{12} = 0.98$ Mbar⁻¹. The dependences on stress \parallel [100] and [110] are shown for a 2DES in the (001) plane in figure 2. Two models have been developed for treating the interaction theoretically. That of Kelly and Falicov [3, 4] was based on charge density waves (CDW) resulting from phonon-mediated intervalley electron exchange interaction. Takada and Ando [5] and Stern and Howard [6] considered the interaction without involving phonons. The CDW model was initially supported by the observation of a single cyclotron resonance (CR) which varied in position continuously between the two masses of the valleys when varying their relative populations. According to the Takada–Ando model the observation of two cyclotron resonances was expected. Results of Shubnikov–de Haas (SdH) experiments could never be uniquely interpreted

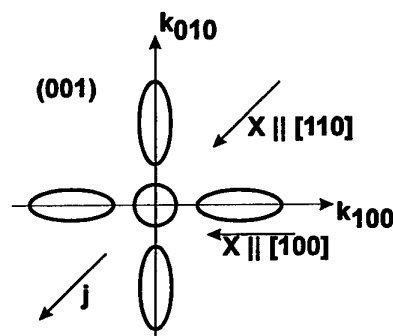


Figure 1. Energy contours of the conduction band in Si MOSFETs and definition of directions.

in favour of one of the two models. Eisele *et al* [7] preferred the CDW interpretation, Englert *et al* [8] doubted that the electron–phonon interaction was strong enough for the occurrence of CDWs. Subsequent CR experiments [9, 10] revealed the existence of two cyclotron masses under conditions where two valleys were expected to be populated. This agreed with the Takada–Ando model. However, in the SdH experiments also performed by Gesch *et al* [11] a constant SdH period, i.e. no indication for a two-valley regime, was observed when varying the stress. Thus, the understanding of the interacting 2D electron system in silicon, the prototype many-valley semiconductor, was far from being satisfactory.

In the present work Si MOSFETs of (001) orientation has been studied under uniaxial stress using high-magnetic field experiments, i.e. evaluating the SdH effect and the quantum Hall effect (QHE). The results allow a phase diagram to be established of the population of the valleys involved. Our results confirm the model of Takada and Ando and clear the discrepancy between the SdH and CR data mentioned above.

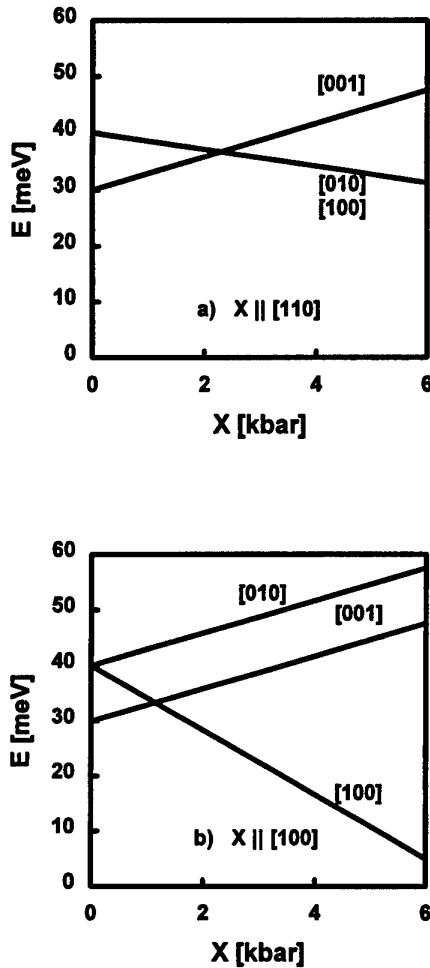


Figure 2. Energetic positions of the valleys in a (001) Si MOSFET for different stress directions without electron-electron interaction: (a) $X \parallel [110]$; (b) $X \parallel [100]$.

2. Experiments and samples

Silicon (001) MOSFETs with standard Hall bar geometry were used for the experiments. Their length and width were 800 and 80 μm , respectively. Two potential probes on each side were spaced by 160 μm . The thickness of the gate oxide was 680 nm. At $T = 2.1$ K the maximum electron mobility was $2 \times 10^4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ (at $V_G = 10$ V). The long axis of the MOSFET was oriented in the [110] direction. The long axis of the sample, i.e. the stress direction, was either [100] or [110]. The total size of the sample was $6 \times 2 \times 0.5 \text{ mm}^3$ ([100] sample) and $6 \times 1 \times 0.5 \text{ mm}^3$ ([110] sample). For the electrical contacts 25 μm diameter gold wires were bonded to the aluminium pads of the MOSFET.

For the magnetotransport experiments under uniaxial stress the samples had to be mounted in a holder where the stress was applied perpendicular to the magnetic field. The force was applied by stainless steel strings via pulleys. The magnetic field of the superconducting magnet was in the vertical direction perpendicular to the (001) plane of the sample. Details of the stress apparatus can be found

in [12]. For applying the force uniformly and for isolating the sample electrically, Mylar foils were positioned between the sample and the pistons. The maximum stress which could be produced was 2.5 and 6 kbar in the samples with $X \parallel [100]$ and [110] respectively.

All the experiments were performed at $T = 2.1$ K with the sample immersed in liquid helium. Gate voltages up to 200 V were applied by a ramp generator. The value of the d.c. drain current was 2 μA , low enough for avoiding distortions of the QHE plateaus. The potential probes were connected to the measuring system (scanner, digital multimeter, computer) via isolation amplifiers with input impedances of 100 G Ω . The quantities deduced were the longitudinal resistivity ρ_{xx} and the Hall resistivity ρ_{xy} . The results reported in this paper were obtained at $B = 10.7$ T. At this field only the lowest subband of the [001] valley is populated at zero stress. The threshold voltage for the onset of conduction was 8 V and remained constant when applying stress and/or magnetic field.

3. Experimental results and interpretation

As examples of our experimental data a few selected $\rho_{xx}(V_G)$ traces with the stress $X \parallel [110]$ as the parameter are shown in figure 3. Above high V_G values the SdH minima clearly shift to higher V_G values with increasing stress above $X \approx 0.6$ kbar. At low V_G values the structure of the ρ_{xx} curves is more complex. The minima of ρ_{xx} correspond to integer filling factors of Landau levels or Landau sublevels (spin and valley splitting taken into account). The actual values of the filling factors are obtained from the numbering of the minima and additionally from the ρ_{xy} values of the corresponding QHE plateaus. The values of the stress, where the minima start to move, vary weakly with the filling factor. Also the periods of the SdH oscillations on the V_G scale vary with increasing stress. At the highest stress the period is 10% larger than at zero stress.

The shifts of the minima at high V_G values are accompanied by the appearance of an additional set of oscillations at low V_G values and high stress. The identification of the minima by their corresponding filling factors is the basis for the construction of the phase diagram for the population of the different valleys. For this procedure it is assumed that the oxide capacitance and the threshold voltage do not change with stress. This is confirmed by the following experimental observations: (i) the low-field Hall effect is independent of the stress; (ii) the minima with the same values of the filling factors appear at the same V_G values indicating no change of the carrier density; (iii) the measured threshold voltage is independent of the stress. The set of SdH oscillations observable at low stress values is attributed to the circular valley and the additional set at low V_G and high X to the elliptical valleys. In the phase diagrams of figure 4 and figure 5 the population of the circular valley, n_{001}/n_S , is shown as a function of n_S and X . The RPA results of Takada and Ando are in excellent qualitative agreement with our experimental data. The agreement could be made more quantitative (observe the different scales in the

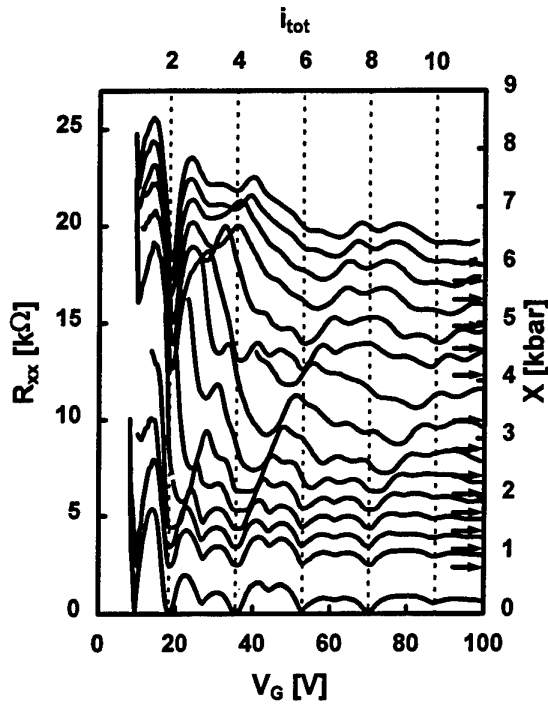


Figure 3. Magnetoresistance as a function of gate voltage with the stress in the [110] direction as a parameter. The curves are shifted proportional to the applied stress. The arrows on the right-hand scale indicate the values of the stress and the zeros of the shifted resistance scales.

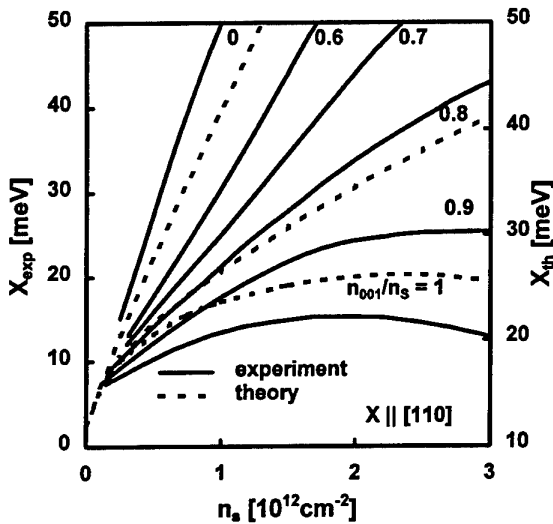


Figure 4. Phase diagram for the (001) MOSFET with uniaxial stress in the [110] direction. The broken curves show the theoretical results after Takada and Ando (private communication) for $n_{001}/n_s = 0, 0.8$ and 1 .

diagrams) if a lower depletion charge ($n_{dep} = 1 \times 10^{11} \text{ cm}^{-2}$ in [5]) was used in the calculations.

In the previous publications the experimental results of the magnetoresistance were considered as being in contradiction to the CR measurements [10] and to the

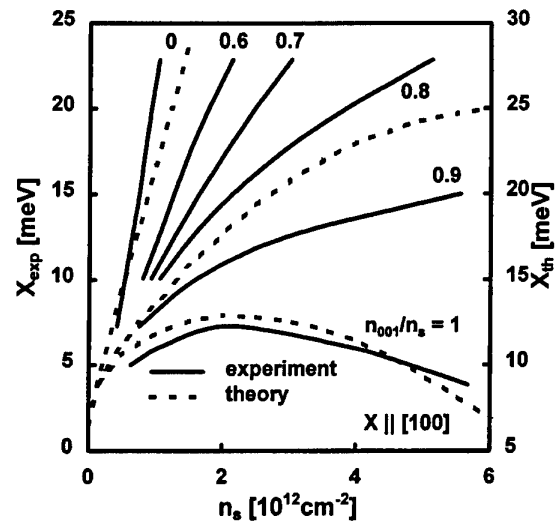


Figure 5. Phase diagram for the (001) MOSFET with uniaxial stress in the [100] direction. The broken curves show the theoretical results after Takada and Ando [5] for $n_{001}/n_s = 0, 0.8$ and 1 .

theoretical results of Takada and Ando [5]. The main feature was a constant SdH period independent of stress without any indication of a second period of another valley. According to [13] at least a stress-dependent period should be observable if the interaction is too weak for the formation of CDWs. A detailed examination of our ρ_{xx} data, however, shows that the SdH period for $V_G > 50 \text{ V}$ (i.e. $n_s > 1.5 \times 10^{12} \text{ cm}^{-2}$) is actually not constant. This could not be observed by Englert *et al* [14] because of lower sample quality and weaker SdH oscillations.

The excellent agreement of our data with the theory of Takada and Ando shows that a CDW-based interpretation is not applicable to the SdH data. Our interpretation is also consistent with the CR measurements by Stallhofer *et al* [9] where two CR masses were observed in the regime which we identified as that of the mixed phase.

In our interpretation we used a comparison of experiments in high magnetic fields and of a $B = 0$ theory. This procedure can be justified by assuming broadened density of states of the Landau levels (valley-split and spin-split sublevels). This broadening leads to an overlap of neighbouring levels. The overlap can be considered as being strong enough [13] to apply the $B = 0$ density of states as a good approximation in high fields. The observation of pronounced minima of R_{xx} should not be seen as contradictory since they are essentially caused by localization in the tails of the density of states of the overlapping sublevels.

Acknowledgment

The authors would like to thank Dr G Dorda for valuable discussions and the fabrication of the MOSFETs.

References

- [1] Herring C and Vogt E 1956 *Phys. Rev.* **101** 944
- [2] Ramdas A K and Rodriguez S 1981 *Rep. Prog. Phys.* **44** 1297
- [3] Kelly M J and Falicov L M 1976 *Phys. Rev. Lett.* **37** 1021
- [4] Kelly M J and Falicov L M 1977 *Phys. Rev. B* **15** 1974
- [5] Takada Y and Ando T 1978 *J. Phys. Soc. Japan* **44** 905
- [6] Stern F and Howard W E 1967 *Phys. Rev.* **163** 816
- [7] Eisele I, Gesch H and Dorda G 1977 *Solid State Commun.* **22** 185
- [8] Englert Th, Landwehr G, Pontcharra J and Dorda G 1980 *Surf. Sci.* **98** 427
- [9] Stallhofer P, Kotthaus J P and Abstreiter G 1979 *Solid State Commun.* **32** 655
- [10] Abstreiter G, Stallhofer P and Kotthaus J P 1980 *Surf. Sci.* **98** 413
- [11] Gesch H, Dorda G, Stallhofer P and Kotthaus J P 1979 *Solid State Commun.* **32** 543
- [12] Kuchar F, Veigl K and Fantner E J 1979 *Rev. Sci. Instrum.* **50** 245
- [13] Ando T, Fowler A B and Stern F 1982 *Rev. Mod. Phys.* **54** 437
- [14] Englert Th, Landwehr G, von Klitzing K, Dorda G and Gesch H 1978 *Phys. Rev. B* **18** 794

SECTION 7. MESOSCOPICS

- 1613 Electrons in mesoscopically inhomogeneous magnetic fields**
P D Ye, D Weiss, R R Gerhardt, G Lütjering, K von Klitzing and H Nickel
- 1618 Universal conductance fluctuations in submicron wires of $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$**
T Dietl, J Jaroszyński, G Grabecki, J Wróbel, M Sawicki, T Skośkiewicz, E Kamińska, A Piotrowska, G Karczewski, T Wojtowicz and J Kossut
- 1624 Metal-insulator transition in Sb-doped short-period Si/SiGe superlattices**
G Brunthaler, T Dietl, M Sawicki, G Stöger, J Jaroszynski, A Prinz, F Schäffler and G Bauer
- 1630 A study of the interacting electron system in stressed (001) Si MOSFETs**
J Lutz and F Kuchar

- 1534 **Towards the optimization of the growth and properties of modulated electron systems on GaAs vicinal surfaces**
B Etienne, F Laruelle, Z Wang, L Sfaxi, F Lelarge, F Petit, T Mélin and A Cavanna

SECTION 4. MAGNETOTRANSPORT AND QUANTUM HALL EFFECT

- 1539 **Correlated bilayer electron states**
M Shayegan, H C Manoharan, Y W Suen, T S Lay and M B Santos
- 1546 **Breakdown of the integer quantum Hall effect at high currents in GaAs/AlGaAs heterostructures**
S Kawaji

SECTION 5. NANOLITHOGRAPHY, NANOFABRICATION AND NANOPROBES

- 1552 **Nanolithography**
D K Ferry, M Khoury, D P Pivin Jr, K M Connolly, T K Whidden, M N Kozicki and D R Allee
- 1558 **Proximal probe-based fabrication of nanostructures**
P M Campbell and E S Snow
- 1563 **STM investigation and manipulation of C₆₀ molecules adsorbed on an Si(111) surface**
A W Dunn, P Moriarty, M D Upward, A Nogaret and P H Beton
- 1569 **High spatial resolution spectroscopy of single semiconductor nanostructures**
T D Harris, D Gershoni, L Pfeiffer, M Nirmal, J K Trautman and J J Macklin
- 1575 **Growth and transport properties of thin Bi films on InP(110)**
B G Briner, R M Feenstra, T P Chin and J M Woodall

SECTION 6. SUPERLATTICES AND MULTIPLE QUANTUM WELLS

- 1582 **Landau subbands generated by a lateral electrostatic superlattice—chasing the Hofstadter butterfly**
T Schlösser, K Ensslin, J P Kotthaus and M Holland
- 1586 **Low-dimensional systems in ultra-high magnetic fields: magnetic-field-induced type I to type II transitions in short-period semiconductor superlattices**
N Miura, Y Shimamoto, Y Imanaka, H Arimoto, H Nojiri, H Kunitatsu, K Uchida, T Fukuda, K Yamanaka, H Momose, N Mori and C Hamaguchi
- 1591 **Intersubband dynamics of asymmetric quantum wells studied by THz 'optical rectification'**
K Unterrainer, J N Heyman, K Craig, B Galdrikian, M S Sherwin, K Campman, P F Hopkins and A C Gossard
- 1596 **Virtual states, dynamic localization, absolute negative conductance and stimulated multiphoton emission in semiconductor superlattices**
B J Keay, C Aversa, S Zeuner, S J Allen Jr, K L Campman, K D Maranowski, A C Gossard, U Bhattacharya and M J W Rodwell
- 1601 **Bandgap engineering of electro-optical modulators and switches**
P Voisin
- 1608 **Effect of donor impurities on far-infrared magnetospectroscopy of electrons in quasi-two-dimensional systems**
Z X Jiang, S R Ryu and B D McCombe

Continued

SEMICONDUCTOR SCIENCE & TECHNOLOGY

Volume 11

Number 11S

November 1996

**9th Winterschool on New Developments in Solid State Physics
Mauterndorf, Austria, 19–23 February 1996**

SECTION 1. COMPOSITE FERMIONS AND SKYRMIONS

- 1473 Examining composite fermions with surface acoustic waves**
R L Willett and L N Pfeiffer
- 1477 The dependence of the Composite Fermion effective mass on carrier density and Zeeman energy**
R J Nicholas, D R Leadley, M S Daly, M van der Burgt, P Gee, J Singleton, D K Maude, J C Portal, J J Harris and C T Foxon
- 1482 Evidence for quasi-classical transport of composite fermions in an inhomogeneous effective magnetic field**
J H Smet, R Fleischmann, D Weiss, R Ketzmerick, R H Blick, G Lütjering, K von Klitzing, T Geisel and G Weimann
- 1488 Optically pumped nuclear magnetic resonance in the quantum Hall regimes**
S E Barrett, G Dabbagh, L N Pfeiffer, K W West and R Tycko

SECTION 2. SINGLE-ELECTRON PHENOMENA

- 1493 Single-electron transistors with a self-assembled quantum dot**
M Dilger, R J Haug, K Eberl and K von Klitzing
- 1498 Non-invasive detection of single-electron processes**
M Field, C G Smith, M Pepper, K M Brown, E H Linfield, M P Grimshaw, D A Ritchie and G A C Jones
- 1502 A possible road to 77 K single-electron devices**
S Altmeyer, F Kühnel, B Spangenberg and H Kurz

SECTION 3. QUANTUM DOTS

- 1506 Tunnelling through quantum dots**
R H Blick, T Schmidt, R Haug and K von Klitzing
- 1512 Photon-assisted tunnelling through a quantum dot**
T H Oosterkamp, L P Kouwenhoven, A E A Koolen, N C van der Vaart and C J P M Harmans
- 1516 Electronic structure and optical properties of self-assembled quantum dots**
P Hawrylak and A Wojs
- 1521 Growth and characterization of self-assembled Ge-rich islands on Si**
G Abstreiter, P Schittenhelm, C Engel, E Silveira, A Zrenner, D Meertens and W Jäger
- 1529 Optical studies of free-standing single InGaAs/GaAs quantum dots**
A Forchel, R Steffen, T Koch, M Michel, M Albrecht and T L Reinecke

Continued on inside back cover

Bibliographic codes CODEN: SSTEET 11 (11S) 1473–1634 (1996) ISSN: 0268-1242